НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»

Інститут телекомунікаційних систем

Кафедра Телекомунікаційних систем

«На правах рукопису»
УДК 621.39

«До захисту допущено»
Завідувач кафедри
_____ Л.О. Уривський
«___»_____20__ р.

Магістерська дисертація

на здобуття ступеня магістра
зі спеціальності 172 Телекомунікації та радіотехніка

на тему: «Дослідження можливостей надання послуг в віртуальному середовищі телекомунікаційних мереж NGN на базі IMS»

Виконав (-ла):
студент (-ка) ІІ курсу, групи ТС-91мп
Гогу Олександр Петрович _____

Керівник:
доцент, кандидат технічних наук
Гаттуров В.К. _____

Рецензент:
доцент, кандидат технічних наук
Мазор С.Ю. _____

Засвідчую, що у цій магістерській дисертації немає запозичень з праць інших авторів без відповідних посилань.
Студент (-ка) _____

Київ – 2020 року

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Інститут телекомунікаційних систем
Кафедра Телекомунікаційних систем

Рівень вищої освіти – другий (магістерський) за освітньо-професійною програмою
Спеціальність (спеціалізація) – 172 «Телекомунікації та радіотехніка» (172.3620.1 «Телекомунікаційні системи та мережі»)

ЗАТВЕРДЖУЮ
Завідувач кафедри
_____ Л.О. Уривський
«___»_____20__ р.

ЗАВДАННЯ
на магістерську дисертацію студенту

Гогу Олександру Петровичу

1. Тема дисертації «Дослідження можливостей надання послуг в віртуальному середовищі телекомунікаційних мереж NGN на базі IMS», науковий керівник дисертації Гаттуров Віктор Кавіч, доцент, кандидат технічних наук, затверджені наказом по університету від «___»_____ 20__ р. №_____

2. Термін подання студентом дисертації 13 грудня 2020 р.

3. Об’єкт дослідження IMS на базі програмно-конфігурованих мереж та віртуалізації мережевих функцій.

4. Предмет дослідження можливості підсистеми IMS для надання послуг у віртуальному середовищі.

5. Перелік завдань, які потрібно розробити

- дослідження інноваційних змін в інфокомунікаційних мережах при їх еволюційному переході до концепції мультимедійної IP підсистеми;
- огляд технологій програмно-конфігурованих мереж та віртуалізації мережевих функцій;
- моделювання кластеру контролерів програмно-конфігурованої мережі для дослідження ефективності її функціонування;
- розробка моделі ідентифікації та приорітезації трафіку Інтернету Речей на основі сегментації ресурсів, та моделювання процесу надходження даних до сегментів запропонованої мережі.

6. Орієнтовний перелік графічного (ілюстративного) матеріалу
Плакат №1 «Тема, мета, актуальність, об’єкт, предмет, завдання дослідження»

Плакат №2
Плакат №3
Плакат №4
Плакат №5
Плакат №6
Плакат №7 «Висновки»

7. Орієнтовний перелік публікацій

Гогу О.П., Гаттуров В.К. ДОСЛІДЖЕННЯ МОЖЛИВОСТЕЙ НАДАННЯ ПОСЛУГ В ВІРТУАЛЬНОМУ СЕРЕДОВИЩІ ТЕЛЕКОМУНІКАЦІЙНИХ МЕРЕЖ NGN НА БАЗІ IMS. ПТ-2020. - К.: КПІ ім. Ігоря Сікорського, 2020.

8. Дата видачі завдання 1 вересня 2019 р.

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів магістерської дисертації	Примітка
1	Огляд літературних джерел по тематиці роботи	01.09.19 - 01.11.19	
2	Формування мети та наукових завдань дослідження	01.11.19 - 01.12.19	
3	Формування структури першого та другого розділу наукової роботи	01.01.20 - 25.03.20	
4	Підготовка матеріалів для доповіді на конференції	25.03.20 - 20.04.20	
5	Формування структури третього розділу наукової роботи	20.04.20 - 30.05.20	
6	Проведення досліджень для вирішення наукових завдань	30.05.20 - 01.08.20	
7	Формування структури четвертого розділу наукової роботи	01.08.20 - 30.08.20	
8	Наповнення чотирьох розділів наукової роботи, висновків	01.09.20 - 05.12.20	
9	Оформлення плакатів та пояснювальної записки дипломної роботи	05.12.20 - 17.12.20	
10	Підготовка до захисту дипломної роботи	17.12.20 - 22.12.20	

Студент Гогу О.П.

Науковий керівник дисертації Гаттуров В.К.

РЕФЕРАТ

Темою магістерської дисертації є дослідження можливостей надання послуг в віртуальному середовищі телекомунікаційних мереж NGN на базі IMS

Робота містить 108 сторінок, зокрема 31 рисунок, 3 таблиці та 35 джерел інформації.

Актуальність теми обумовлена розвитком та застосуванням мереж наступного покоління (NGN) та їх конвергенцією з мобільними мережами 5G і як наслідок, утворення єдиної мережі з реалізацією мультимедійної IP підсистеми, окремі мережеві функції якої реалізовані віртуально.

Мета даної роботи полягає у підтвердженні можливостей надання нових послуг абонентам інфокомунікаційної мережі, при реалізації архітектури IMS.

Об’єкт дослідження IMS на базі програмно-конфігурованих мереж та віртуалізації мережевих функцій.

Предмет дослідження можливості підсистеми IMS для надання послуг у віртуальному середовищі.

В даній роботі з’ясовується роль мультимедійної IP підсистеми та її зв’язок з мобільними мережами 5G. Проводиться імітаційне моделювання кластеру контролерів ПКМ для оцінки ефективності їх функціонування. На основі проведеного моделювання розроблюється модель мережі зв’язку на основі сегментації ресурсів.

КОНЦЕПЦІЯ, АРХІТЕКТУРА, IP MULTIMEDIA SUBSYSTEM, 5G, NETWORK FUNCTION VIRTUALIZATION, SOFTWARE DEFINED NETWORK, ПОСЛУГА

ABSTRACT

The relevance of the topic is due to the development and application of next generation networks (NGN) and their convergence with 5G mobile networks and, as a consequence, the formation of a single network with the implementation of multimedia IP subsystem, some network functions are implemented virtually.

The purpose of this work is to confirm the possibility of providing new services to subscribers of the infocommunication network, in the implementation of the IMS architecture.

The object of IMS research on the basis of software defined networks and network functions virtualization.

The subject of research is the possibility of the IMS subsystem to provide services in a virtual environment.

This paper examines the role of the multimedia IP subsystem and its connection to 5G mobile networks. Simulation modeling of a cluster of SDN controllers is carried out for an estimation of efficiency of their functioning. Based on the simulation model developed communication network based on the segmentation of resources.

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ

3GPP	3rd Generation Partnership Project
5G	5th Generation Mobile Network
AAA	Authentication, Authorization and Accounting
AMF	Access and Mobility Management Function
AS	Application Server
BICC	Bearer independent call control protocol
BSS	Business Support System
CAMEL	Customized Applications for Mobile Network Enhanced Logic
CDMA	Code Division Multiple Access
CDR	Call Detail Records
CGF	Charging Gateway
COTS	Commercial Off The Shelf
CS	Call Server
DC	Data center
DDC	Dynamic and Distributed Clustering
DHCP	Dynamic Host Configuration Protocol
DTLS	Datagram Transport Layer Security
ETSI	European Telecommunications Standards Institute
GGSN	Gateway GPRS Service Node
GPRS	General Packet Radio Service
GSM	Global System for Mobile Communications
HFC	Hybrid Fiber Coax
HLR	Home Location Register

HSS	Home Subscriber Server
HTTP	Hyper Text Transfer Protocol
I-CSCF	Interrogating Call Session Control Function
IMS	IP Multimedia Subsystem
IoT	Internet of Things
IP	Internet Protocol
ISDN	Integrated Services Digital Network
ISUP	ISDN User Part
LINP	Logically Isolated Network Partition
LTE	Long Term Evolution
MA	Management Application
MAC	Media Access Control
MANO	NFV Management and Orchestration
MGCF	Media Gateway Control Function;
MGW	Media Gateway
MQTT	Message Queue Telemetry Transport
MRF	Media Resource Function
MRFC	Media Resource Function Controller
MRFP	Media Resource Function Processor
MSC	Mobile Switching Center
NF	Network Function
NFV	Network Function Virtualization
NFVI	NFV Infrastructure
NFVO	NFV Ochestrator
NGN	Next Generation Network

ONF	Open Networking Foundation
OSA	Open Service Access
OSS	Operation Support System
P-CSCF	Proxy Call Session Control Function
PDF	Policy Decision Function
PSTN	Public Switched Telephone Network
QoS	Quality of service
RAN	Radio Access Network
RTP	Real-time Transport Protocol
S-CSCF	Serving Call Session Control Function
SDN	Software Defined Network
SGW	Signaling Gateway
SIP	Session Initiation Protocol
SLF	Subscriber Locator Function
SMF	Session Management Function
SSF	Service Switching Function
SSL	Secure Sockets Layer
TCP	Transmission Control Protocol
TDM	Time Division Multiplexing
TLS	Transport Layer Security
UDP	User Datagram Protocol
UPF	User Plane Function
URI	Uniform Resource Identifier
VNF	Virtual Network Function
Vo5G	Voice over 5G

VoIP	Voice Over IP
VoLTE	Voice over LTE
VR	Virtual resource
WiFi	Wireless Fidelity
xDSL	Digital Subscriber Line
АТС	Автоматична телефонна станція
ГК	Головний контролер
ІР	Інтернет Речей
РК	Розподілений контролер
СМО	Система масового обслуговування
ТмЗК	Телефонні мережі загального користування

ВСТУП

Протягом останнього десятиліття телекомунікаційна індустрія спостерігала тенденцію зростання кількості мереж фіксованого та мобільного зв’язку. Це привело до того, що між провайдерами йде активна боротьба за користувачів, так як ринок послуг мобільного зв'язку поступово переходить в стан насичення, темпи зростання абонентської бази операторів значно знижуються, тобто подальше залучення абонентів можливе тільки за рахунок переходу абонентів від одного оператора до іншого, але не за рахунок підключення нових. Виходить так, що прибуток оператора зв’язку безпосередньо залежить від якості послуг, які вони надають, а це в свою чергу свідчить, що провайдерам потрібно шукати нові шляхи вирішення даного питання.

Тому на ринку відбуваються активні процеси злиття і поглинання. Конвергенція фіксованих і мобільних технологій представляє інтеграцію технологій та послуг дротового та бездротового зв'язку для створення єдиного телекомунікаційного мережевого середовища. Така ідея обіцяє подолати деякі фізичні бар'єри, які зараз заважають провайдерам телекомунікаційних послуг досягти максимуму своїх потенційних клієнтів.

Концепція IMS (IP Multimedia Subsystem) обіцяє зручне для оператора середовище для пакетних дзвінків та послуг у режимі реального часу, які не тільки дозволять зберегти традиційний контроль оператора над сигналізацією користувача та оплатою на основі використання, але й принесуть новий дохід з легким розгортанням широко спектру послуг користувачу. Крім того, IMS архітектура проектувалася як незалежна від технології мережі доступу, що дозволяє отримувати доступ до IMS послуг з будь-якого терміналу по будь-якої зручної технології доступу.

Очікується, що мережі зв'язку п'ятого покоління (5G) забезпечуватимуть постачання наскрізних послуг з гарантованою якістю обслуговування для величезної кількості підключених пристроїв Інтернету Речей, підтримку

різноманітних варіантів використання і додатків (у тому числі розумні будинки, промислова автоматизація, розумні транспортні системи і системи електронної охорони здоров'я). В свою чергу запуск будь-якого нового мережевого сервісу передбачає витрати на обслуговування, ремонт, заміну та купівлю нового обладнання, що вимагає місця в апаратних кімнатах, нових джерел живлення. Універсальним виходом з даної ситуації стала віртуалізація мережевих функцій NFV, пов'язана з концепцією програмно-конфігурованих мереж SDN. Консолідування цих технологій допоможе розмити границі між вендорами, розв’язати проблему росту пакетного трафіку та напряму допоможе операторам раціоналізувати витрати на обслуговування і управління мережею. Також поєднання цих технологій дозволяє реалізувати основну бізнес ідею мереж 5G – сегментування. Яка, в свою чергу, забезпечує гнучкість мережі, розбиваючи одну фізичну мережу на кілька шарів, кожен з яких має власні налаштування, адаптовані під певну послугу. Таким чином забезпечується ефективність і гнучкість майбутніх сервісів.

У даній роботі на базі проведених експериментальних досліджень буде розроблена модель мережі оператора зв’язку, яка традиційно розділена на 3 сегменти: рівень доступу, рівень агрегації, рівень ядра мережі. При цьому кожним з цих сегментів буде управляти SDN-контролер. В рамках розробки моделі буде проведено імітаційне моделювання сегмента запропонованої мережі.

1 СТРУКТУРА ТА КОНЦЕПЦІЯ IMS В СУЧАСНОМУ ТЕЛЕКОМУНІКАЦІЙНОМУ СВІТІ

1.1 Перехід до мереж нового покоління

Для того, щоб зрозуміти причини появи такої мережі, як NGN (Next Generation Network), потрібно звернутися до історії телекомунікацій, а почнемо ми наші мандри з початку цифровізації.

При переході від аналогових до цифрових АТС (Автоматична телефонна станція), в добавок до традиційних голосових дзвінків, оператори зв'язку отримали можливість надавати користувачам розширені пакети послуг. В геометричній прогресії ріс обсяг передачі різного виду даних, з'являлось безліч локальних і глобальних інформаційних мереж, в тому числі глобальна мережа інтернет, передача даних усередині якої відбувається по IP-протоколу. Однак швидкість передачі даних через інтернет за допомогою звичайного модему, підключеного до ТмЗК (Телефонні мережі загального користування) занадто мала для того, щоб передавати великі обсяги даних, в зв'язку з чим на ринок стали виходити телекомунікаційні компанії, що пропонували користувачам високошвидкісний доступ в інтернет. Але при цьому постає необхідність покупки модему з високою швидкістю передачі, який би забезпечував зв'язок по цифрових каналах зв'язку, а також постає питання сполучення обладнання оператора, який надає високошвидкісний інтернет з обладнанням ТмЗК, яка надає лінії зв'язку.

Створення мережі нового покоління, що представляє собою комплексну мультисервісну мережу, є найкращим вирішенням даних проблем. NGN дозволяє передавати голос, дані, відеоінформацію і інші види послуг по єдиній інфраструктурі мережі. Особливістю цієї мережі є те, що елементи та обладнання (канали, маршрутизатори, комутатори, шлюзи) для передачі і маршрутизації пакетів, з фізичної та логічної точок зору відокремлені від пристроїв і логіки управління послугами та викликами. Логіка, яка використовується в мережі, підтримує всі послуги з принципом комутації

пакетів, починаючи від стандартного телефонного зв'язку і закінчуючи послугами передачі даних, мультимедіа - інформації, широкосмугових і керуючих додатків.

Першим кроком до створення NGN можна рахувати розробку проекту розподіленої моделі мережевого шлюзу. Вона фізично розподілила функції управління викликами і функції підтримки сесії передачі даних, розташувавши їх в різних блоках – управляючих і медіа шлюзах, відповідно. Згодом, дана модель, отримавши ряд доповнень, стала повноцінним мережевим рішенням та здобула назву SoftSwitch. У даній архітектурі управління викликами фізично відокремлена від комутації завдяки поділу на транспортний рівень і рівень управління викликами і сигналізацією. Надання послуг і додатків користувачам також винесено в окремий рівень. Всіма трьома рівнями управляє четвертий - рівень експлуатаційного управління. Мережа на основі SoftSwitch забезпечує комутацію пакетів, а основний протокол для взаємодії - SIP (Session Initiation Protocol), найефективніший і простий IP-протокол, він ефективно використовує інтернет-технології, а також дає можливість впровадити спектр нових послуг і додатків.

1.2 IMS як новий подих для NGN

Створення архітектури NGN відіграло важливу роль в історії телекомунікацій. Але творчий про процес не стоїть на місці, "з ребра" NGN "зліпили" новий підхід до організації різноманітних телекомунікаційних послуг для любого користувача і в будь-якому місці мережі наступного покоління. Цей новий підхід, названий IMS (IP Multimedia Subsystem), виник в результаті еволюції мереж, побудованих на базі технології SoftSwitch, до якої була додана область управління мультимедійними сеансами на базі протоколу SIP та широкий спектр мультимедійних послуг, включаючи, розуміється, двосторонній аудіо та відеозв’язок. Втім, принципово нова концепція IMS в значній мірі виявилася реінкарнацією ранньої ідеї SoftSwitch. Більш того,

спочатку це було просто перенесення архітектури SoftSwitch на мобільні мережі зі збереженням все тих же трьох рівнів:

- User Plane - користувацький рівень, або рівень передачі даних;
- Control Plane - рівень управління;
- Application Plane - рівень додатків;

В концепціях IMS специфікуються вузли мережі, а функції, які зв'язуються один з одним через стандартизовані еталонні точки. [1]

Вона обіцяє зручне для оператора середовище для пакетних дзвінків та послуг у режимі реального часу, які не тільки дозволять зберегти традиційний контроль оператора над сигналізацією користувача та оплатою на основі використання, але й принесуть новий дохід з легким розгортанням широко спектру послуг користувачу.

1.3 Архітектура IMS

Для IMS розроблена багаторівнева архітектура з поділом транспорту для перенесення трафіку і сигнальної мережі IMS для управління сеансами (рисунок 1.1). Таким чином, 3GPP (3rd Generation Partnership Project) при розробці IMS фактично переніс основну ідеологію SoftSwitch на мобільні мережі. Хоча деякі функції не завжди легко віднести до того чи іншого рівня, але такий підхід забезпечує мінімальну залежність між рівнями. [2]

У IMS можна виділити трьох рівневу архітектуру:

- User Plane - рівень передачі даних;
- Control Plane - рівень управління;
- Application Plane - рівень додатків.

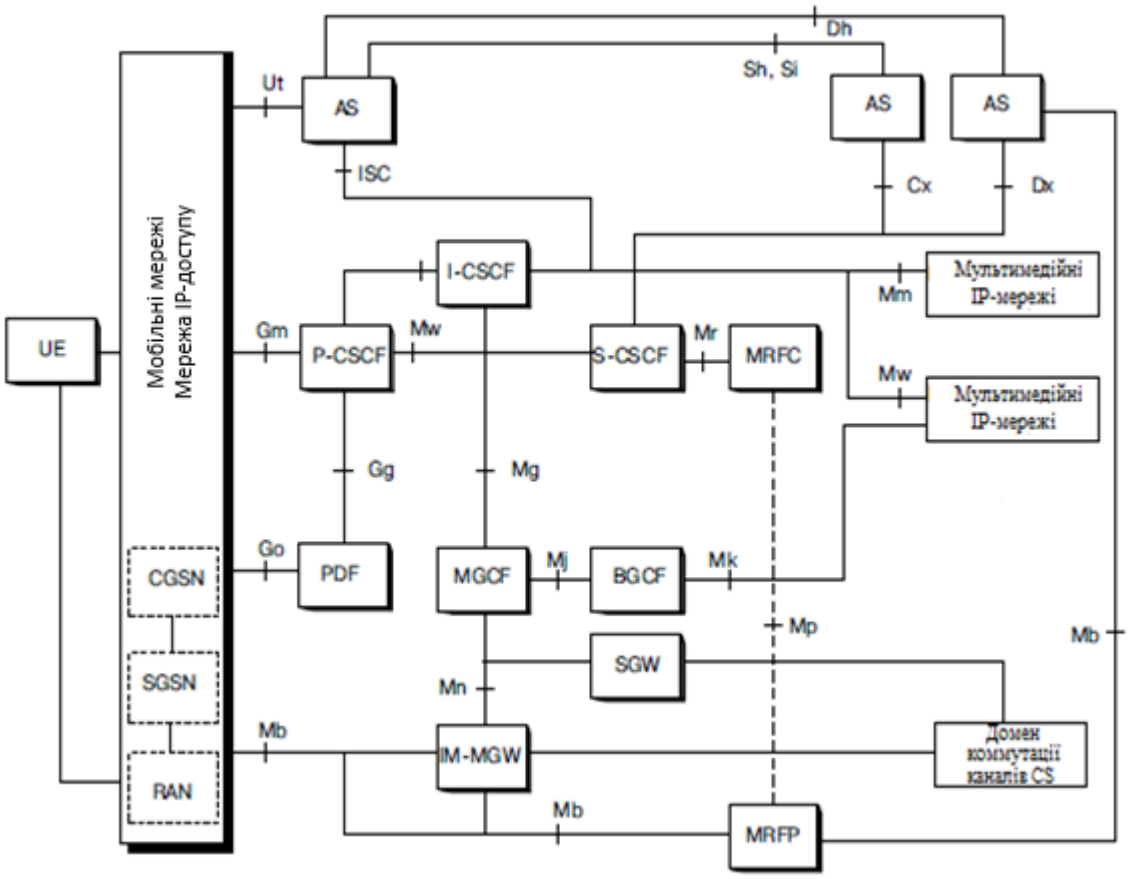

Рисунок 1.1 Архітектура мережі IMS

Перш ніж переходити розгляду елементів IMS-архітектури потрібно розуміти, що організація 3GPP, за прикладом як інтелектуальної мережі, так і SoftSwitch, специфікує не вузли мережі, а їх функції. Це означає, що IMS-архітектура являє собою набір функцій, з'єднаних стандартними інтерфейсами між ними. Розробники мають право як комбінувати кілька функцій в одному фізичному об'єкті так і, навпаки, реалізувати одну функцію розподіленою. Але все ж найчастіше фізичну архітектуру ставлять у відповідність функціональної і реалізують кожну функцію в окремому вузлі. [2]

Різні функції IMS зв'язуються між собою через набір контрольних точок, подібно концепції ISDN (Integrated Services Digital Network). Як і в архітектурі ISDN, еталонні точки в IMS використовуються для ідентифікації та опису інтерфейсів між різними функціональними мережевими елементами. [2]

1.3.1 Функціональні вузли IMS

1) Call Session Control Function (CSCF).

Функція управління сеансами CSCF є центральною частиною системи IMS, являє собою, по суті, SIP-сервер і обробляє SIP сигналізацію в IMS. Існують функції CSCF трьох типів:

- Proxy Call Session Control Function (P-CSCF);
- Interrogating Call Session Control Function (I-CSCF);
- Serving Call Session Control Function (S-CSCF).

Відповідно до концепції IMS, кожна сигнальна подія, яку генерує користувач, спершу направляється до P-CSCF незалежно від самої сигнального події, яка може передбачати такі речі як запит на встановлення сеансу зв'язку, активізацію необхідної функції, виділення мережевого ресурсу, запит обслуговування іншим додатком. Таким чином, функція P-CSCF є першим з ким зустрічається користувач в IMS-ядрі. [1]

Отримавши повідомлення SIP від пристрою користувача, P-CSCF пересилає їх функції до I-CSCF або функції S-CSCF. Функція I-CSCF служить єдиною точкою реєстрації в мережі для доступу до послуг IMS як для місцевих, так і роумінгових користувачів. Як тільки I-CSCF реєструє користувача, S-CSCF вступає до роботи і починає процес управління сеансом зв'язку, забезпечуючи доступ до всіх необхідних служб. [1]

У більшості випадків I-CSCF поводиться як SIP-проксі незважаючи на те, що його основна роль в IMS полягає у визначенні місця надання послуг.

Підкреслимо, що може бути багато I-CSCF в межах мережі постачальника послуг, які відповідальні за наступні функції:

- реєстрація, яка є процесом призначення S-CSCF;
- управління потоком сеансу, що використовується для того, щоб направити SIP запит, який було отримано з іншої мережі, до S-CSCF, або направити запити SIP від користувачів на різні S-CSCF;

- створення даних в формі Call Detail Records (CDR) для білінгу, що генерують дохід від використання ресурсу.

Насправді, S-CSCF виконує роль проксі-сервера SIP, який сеанс за сеансом виконує управління входженням в зв'язок абонентів мережевої служби. Він керує взаємодією з базами даних мережі, такими як домашній абонентський сервер HSS (Home Subscriber Server) для підтримки мобільних користувачів і серверів аутентифікації, авторизації та обліку AAA (Authentication, Authorization and Accounting). У цій ролі S-CSCF також передає сигнальний трафік, що генерується користувачами роумінгу (тобто, користувачами, які під'єднані до мережі в якості гостей) в їхні домашні мережі, де зберігається профіль абонентів і звідки надсилаються підтвердження їх платоспроможності. Як тільки S-CSCF виконав свої обов'язки перевірки користувача, запит може бути переданий відповідним серверам додатків і шлюзів, які потрібні для організації запитуваного сеансу. Якщо запит призначається для іншої мережі, то цілком ймовірно, що доведеться застосувати інші протоколи сигналізації, здійснити перетворення протоколу, організувати взаємодію з іншими типами медіа в інших мережах. [1]

Центральним для роботи IMS є P-CSCF, причому в міру розвитку архітектури IMS його обов'язки трансформувались. Спочатку P-CSCF ніс відповідальність за функцію вибору політики обслуговування PDF (Policy Decision Function), яка накопичує, зберігає, управляє і звертається до політик, щоб прийняти рішення, пов'язані із запитами розподілу ресурсів IP. У міру розвитку архітектури IMS, після низки дискусій PDF була виведена зі сфери P-CSCF, щоб зробити її більш доступною для бездротових LAN (Local Area Network) і інших мереж доступу. Що ж стосується елементів P-CSCF, то в силу їх ролі «воротаря» і того факту, що вони ідеально підходять для збору даних про сесіях, P-CSCF генерують записи білінгу, які можуть бути накопичені і передані централізованій функції Charging Gateway Function (CGF) перед генерацією користувальницьких рахунків. [1]

2) Бази користувачів HSS і SLF

Кожна IMS-мережа містить один або більше серверів користувацьких баз даних HSS. По суті, HSS є централізованим сховищем інформації про абонентів і послуги та являється еволюційним розвитком HLR (Home Location Register) з архітектури мереж GSM (Global System for Mobile Communications). У HSS зберігається вся інформація, яка може знадобитися при встановленні мультимедійного сеансу: інформація про місцезнаходження користувача, інформація для забезпечення захисту (аутентифікація і авторизація), інформація про профіль користувача, про обслуговуючу користувача S-CSCF, про тригерні точки звернення до послуг. На основі записів в базі даних ідентифікується, до яких послуг користувач має доступ, з якою мережею він зараз з'єднаний. HSS підтримує також локалізацію користувача подібно домашньому регістру місцезнаходження HLR (Home Location Register) і тимчасовому регістру місцезнаходження HLR (Home Location Register в мобільних системах попередніх поколінь. Функції, що виконуються HSS, показані в загальному вигляді на рисунку 1.2. [3]

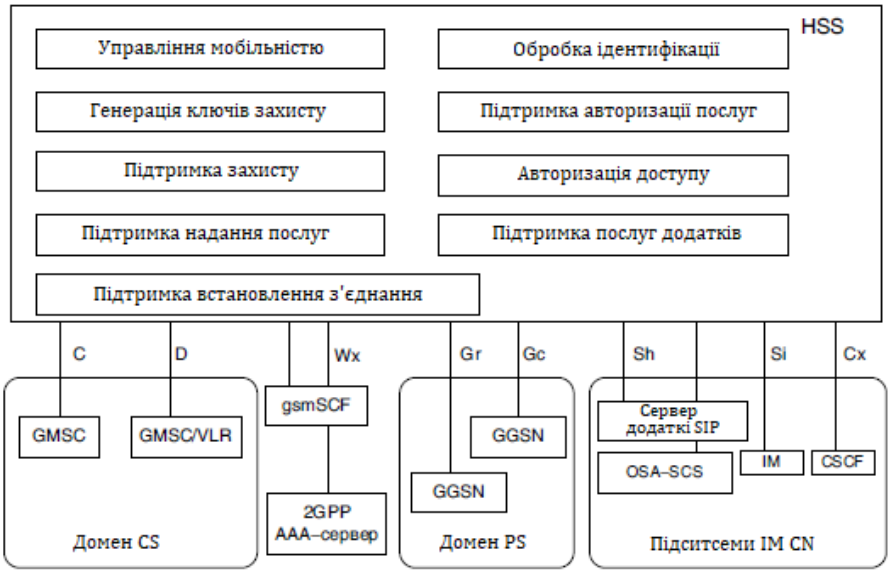

Рисунок 1.2 Логічні функції HSS

Мережа може містити більше одного HSS в тому випадку, якщо кількість абонентів занадто велике, щоб підтримуватися одним HSS. Така мережа, поряд з кількома HSS, повинна буде мати в своєму складі функцію SLF (Subscriber Locator Function), що представляє собою просту базу даних, яка зберігає дані про відповідність інформації HSS адресами користувачів. Вузол, який передав до SLF запит з адресою користувача, отримує від неї відомості про те HSS, який містить інформацію про цього користувача. Як HSS, так і SLF використовують для взаємодії з іншими елементами IMS протокол Diameter. [3]

3) Policy Decision Function (PDF)

Функція PDF іноді інтегрується з P-CSCF, але може бути реалізована окремо. Ця функція відповідає за вироблення політики на підставі інформації про характер сеансу і про переданий трафік (транспортні адреси, ширина смуги і т.д.), отриманої від P-CSCF. На базі цієї інформації PDF приймає рішення про авторизацію запитів від GGSN (Gateway GPRS Service Node) і виробляє повторну авторизацію при зміні параметрів сеансу, а також може заборонити передачу певного трафіку або організацію сеансів деяких типів. [3]

4) Сервери доданків

Сервери додатків AS (Application Server), по суті, не є елементами IMS, а працюють як би поверх неї, надаючи послуги в мережах, побудованих згідно IMS-архітектурі. [3]

Сервери додатків взаємодіють з функцією S-CSCF по протоколу SIP. Основними функціями серверів додатків є обслуговування і модифікація SIP-сеансу, створення SIP-запитів, передача тарифікаційної інформації в центри нарахування плати за послуги. Сервери додатків можуть бути дуже різними, але в IMS прийнято виділяти три типи серверів: SIP AS, OSA-SCS (Open Service Access - Service Capability Server), IM-SSF (IP Multimedia Service Switching Function):

- SIP AS - класичний сервер додатків, що надає мультимедійні послуги на базі протоколу SIP;

- OSA-SCS надає інтерфейс до сервера додатків OSA (Open Service Access) і функціонує як сервер додатків з боку S-CSCF і як інтерфейс між сервером додатків OSA і OSA API (Application Programming Interface) - з іншого боку;
- IM-SSF дозволяє використовувати в IMS послуги CAMEL (Customized Applications for Mobile Network Enhanced Logic), розроблені для GSM мереж, а також дозволяє керуючої функції gsmSCF (GSM Service Control Function) керувати IMS-сеансом. З боку S-CSCF сервер IM-SSF функціонує як сервер додатків, а з іншого боку - як функція SSF (Service Switching Function), що взаємодіє з gsmSCF по протоколу CAP (CAMEL Application Part).

Сервери додатків можуть перебувати або в домашній, або в будь-який інший мережі, з якої у провайдера є сервісну угоду. Але якщо сервер додатків знаходиться у зовнішній мережі, він не може мати інтерфейс з HSS. [3]

5) MRF (Media Resource Function)

Функції медіасерверів в IMS віддані логічному компоненту, відомому як Media Resource Function (MRF), який служить джерелом медіа-інформації в домашній мережі, дозволяє відтворювати мультимедійні оголошення, змішувати медіа-потоки, транскодувати бітові потоки кодеків, отримувати статистичні дані та аналізувати медіа-інформацію. Функції MRF діляться на дві частини:

- MRFC - Media Resource Function Controller
- MRFP - Media Resource Function Processor

MRFC знаходиться на сигнальному рівні і взаємодіє з S-CSCF по протоколу SIP. Використовуючи отримані інструкції, MRFC управляє по протоколу MEGACO (Media Gateway Control Protocol)/H.248 процесором MRFP, що знаходяться на рівні передачі даних, а той виконує всі маніпуляції з медіа-інформацією. Самі MRF завжди знаходяться в домашній мережі. [3]

6) BGCF (Breakout Gateway Control Function)

Breakout Gateway Control Function - це SIP-сервер, здатний виконувати маршрутизацію викликів на основі телефонних номерів. BGCF використовується тільки в тих випадках, коли сеанс ініціюється IMS-терміналом, а адресатом є абонент мережі з комутацією каналів (наприклад, ТмЗК або мобільної мережі 2G). Основними завданнями BGCF є вибір тієї IMS-мережі, в якій повинно відбуватися взаємодія з мережею комутації каналів, або вибір відповідного PSTN (Public Switched Telephone Network)/CS (Chanel switching) шлюзу, якщо ця взаємодія має відбуватися в мережі, де знаходиться сам сервер BGCF. [3]

У першому випадку BGCF переводить сеанс до BGCF обраної мережі, а в другому - до вибраного PSTN/CS шлюзу. [3]

7) Шлюз PSTN/CS

Шлюз PSTN/CS (рисунок 1.3) підтримує взаємодію IMS-мережі з ТМЗК і дозволяє встановлювати з'єднання між користувачами цих мереж.

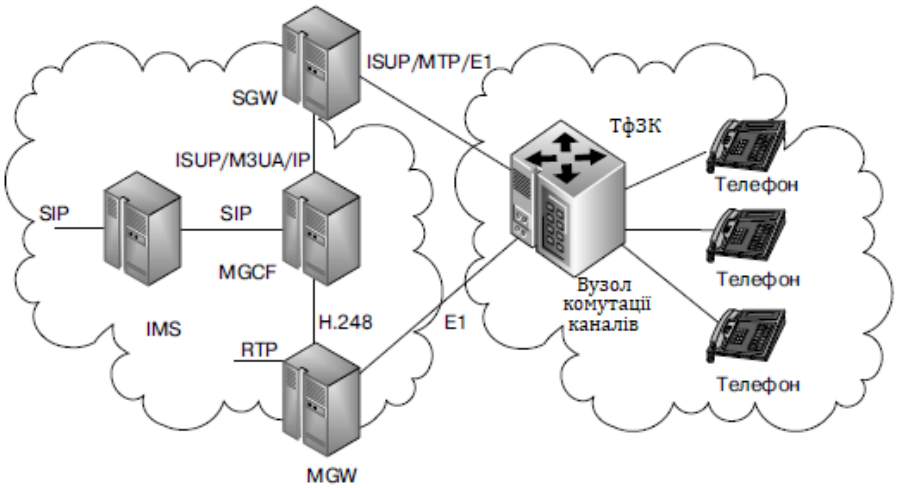

Рисунок 1.3 Шлюз PSTN/CS

Шлюз PSTN/CS має розподілену структуру, характерну для архітектури Softswitch:

- SGW - Signaling Gateway;

- MGCF - Media Gateway Control Function;
- MGW - Media Gateway;

Шлюз сигналізації SGW являється інтерфейсом зв'язку з рівнем сигналізації в мережі комутації каналів, він проводить перетворення нижніх протокольних рівнів систем сигналізації для двостороннього сигнального обміну між мережею IP і мережею ТмЗК. При цьому SGW ніяк не виконує жодних маніпуляцій з повідомленнями прикладного рівня. [3]

Функція управління медіа-шлюзом MGCF – центральна частина розподіленого шлюзу PSTN/CS. Вона перетворює повідомлення ISUP (ISDN User Part) і BICC (Bearer independent call control protocol), які надходять з боку ТМЗК, в повідомлення SIP, які IMS використовує для управління сеансом зв'язку між мережами різних операторів зв’язку (рисунок 1.4). Крім меппінга сигнальних протоколів, MGCF управляє по протоколу MEGACO/H.248 ресурсами медіашлюзу, який бере участь в створенні з'єднання. [3]

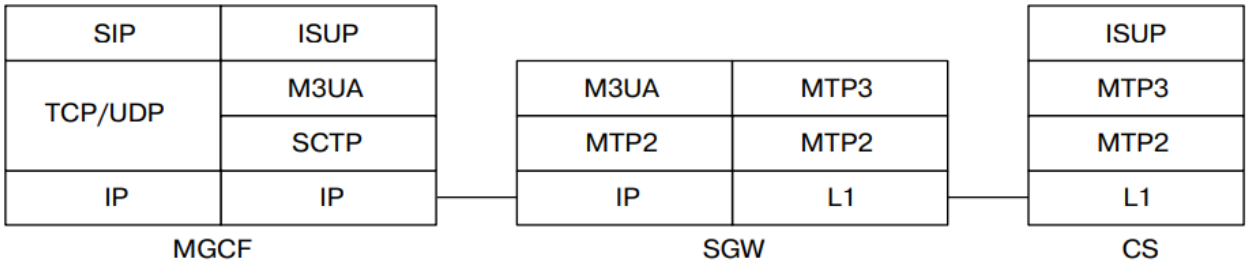

Рисунок 1.4 Управління сеансом зв’язку

Транспортний шлюз MGW виконує функцію з'єднання IP-мережі з мережами комутації каналів на рівні передачі трафіку, виконуючи двостороннє перетворення трафіку користувача, що проходить через границю двох вище описаних мереж. З боку IMS-мережі шлюз MGW веде передачу і прийом даних у вигляді RTP (Real-time Transport Protocol)-пакетів, а з боку мережі з комутацією каналів реалізує стандартний TDM (Time Division Multiplexing)-інтерфейс. Крім того, MGW може здійснювати транскодування інформації, за

умови якщо в IMS і в мережі комутації каналів використовуються різні мовні кодеки (зазвичай в IMS використовується AMR, а в ТМЗК G.711). [3]

1.3.2 Функціональні переваги IMS

Архітектура IMS у світі телекомунікацій, це як Майкл Джордан в баскетболі. На перший погляд звичайна архітектура, але насправді це революція у всіх її аспектах. Ви можете сказати, що сучасні мережі мобільного зв'язку вже і так можуть надавати досить широкий спектр інтернет-послуг та будете правими. Але підхід IMS націлений як на користувача, так і на оператора.

Пояснимо це. Надання різноманітних послуг на базі єдиної пакетної мережі вимагає гнучкої підтримки якості цих послуг. IMS, встановлюючи кожне з'єднання, стежить, щоб користувачам було забезпечено відповідну якість обслуговування. Іншим достатком IMS є ускладнення системи нарахування плати за мультимедійні сеанси зв'язку.

Якщо оператор не приймає до уваги характер трафіку мультимедійного сеансу, він може нарахувати плату за нього тільки дуже поверхневим способом - на підставі обсягу переданих даних. При цьому користувачу невигідно користуватися одними послугами, які створюють великий обсяг трафіку, а оператору невигідно надавати інші, що створюють незначний обсяг трафіку. Якщо виконавець обізнаний про характер переданого трафіку, то він може використовувати в системі нарахування плати більш ефективні бізнес-моделі, що несуть вигоду і йому і користувачам.

У IMS застосований новий підхід до надання послуг, що дозволяє оператору впроваджувати послуги, створені сторонніми розробниками або навіть самим оператором, а не виробниками телекомунікаційного обладнання. Це дозволяє інтегрувати різні послуги і надає широкі можливості персоналізації і збільшення кількості послуг. До впровадження IMS, в мережах використовувалися так звані «вертикальні сервісні платформи», підхід IMS

передбачає горизонтальну архітектуру (Рисунок 1.5), що дозволяє оператору просто і економічно впроваджувати нові персоналізовані послуги, причому користувачі можуть в рамках одного сеансу зв'язку отримати доступ до різних послуг. Перерахуємо деякі функціональні можливості IMS.

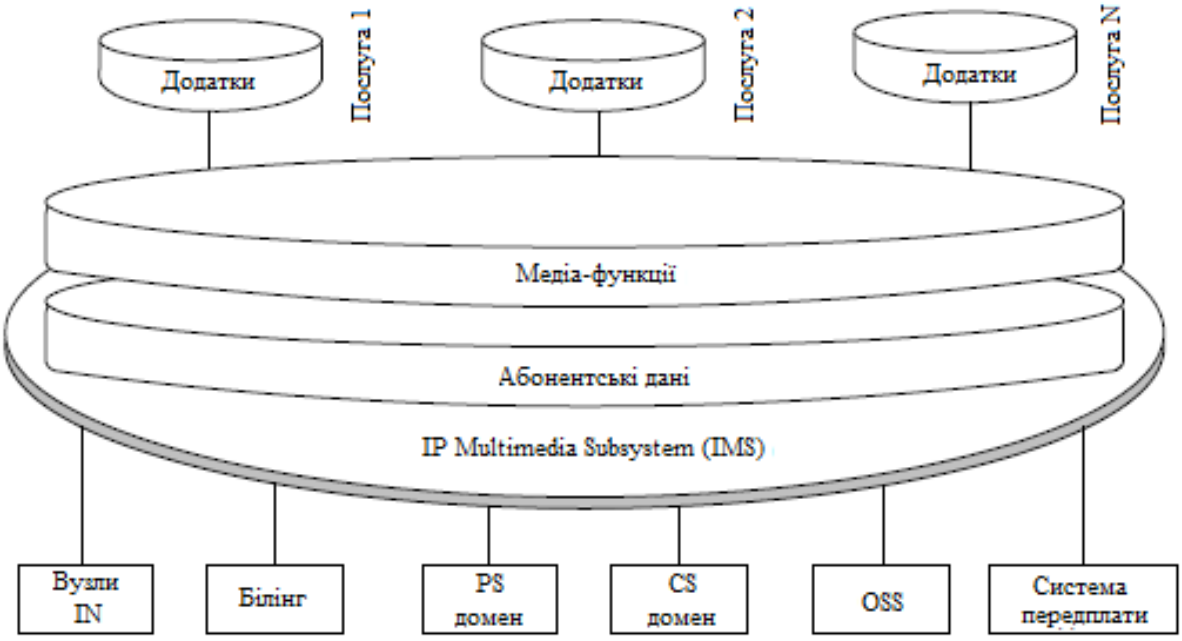

Рисунок 1.5 Горизонтальні сервісні платформи

1) Мультимедійні IP-сеанси

IMS може надавати широкий спектр послуг, але одна з них безумовно зберігає провідну роль - двостороння аудіо відео зв'язок. Для цього архітектура IMS повинна підтримувати сеанси мультимедійного зв'язку в IP-мережах, причому такий зв'язок має бути доступна користувачам як в домашній, так і в гостьовій мережі. Мультимедійна зв'язок була стандартизована вже в ранніх документах 3GPP, ще до появи IMS, але надавалася тільки доменом комутації каналів. [2]

2) Якість обслуговування

Підтримка QoS (Quality of service) є фундаментальною вимогою до IMS. При організації сеансу користувача устаткування сповіщає IMS про свої можливості і про свої вимоги до QoS. За допомогою протоколу SIP можливо врахувати такі параметри, як тип і напрямок передачі даних, бітова швидкість,

розмір пакетів, використання RTP, необхідна ширина смуги пропускання. IMS дозволяє управляти якістю зв'язку, яке отримає той чи інший користувач, і таким чином диференціювати користувачів і послуги. [2]

3) Взаємодія з іншими мережами

Функція підтримки взаємодії мережі IMS з мережею інтернет очевидна, тому що завдяки спільним протоколам користувачі IMS можуть встановлювати мультимедійні сеанси зв'язку з різними службами глобальної мережі. Так як перехід до тотального використання IMS буде поступовим і більш-менш тривалим, IMS повинна також мати можливість взаємодії з мережами попередніх поколінь – стаціонарними (ТмЗК) і мобільними мережами (2G, 3G) років з комутацією каналів і комутацією пакетів. Зрозуміло, що взаємодія з мережами комутації каналів не має довгострокової перспективи, але вона все ж необхідна до повного виведення даної технології з сучасних телекомунікацій. [2]

4) Інваріантність доступу

Оскільки IP є базовим протоколом доступу в IMS, користувач може з'єднатися з мережею безліччю способів. Пристрої з можливостями IMS можуть безпосередньо зареєструватися в мережі IMS незалежно від тієї мережі, з якої вони з'єднуються. [2]

Технології фіксованого доступу, такі як Ethernet LAN, xDSL (Digital Subscriber Line), HFC (Hybrid Fiber Coax), модеми xDSL, а також бездротовий доступ GSM, GPRS (General Packet Radio Service), CDMA (Code Division Multiple Access), WiFi (Wireless Fidelity) підтримуються IMS. Точно так же традиційні телефони ТМЗК і деякі системи VoIP (Voice Over IP) з'єднуються з мережею IMS за допомогою відповідних шлюзів. Таким чином, як і будь-яка IP-мережа, IMS інваріантна щодо протоколів нижніх рівнів і технологій доступу. Вже з шостого релізу версії IMS, функції доступу були відокремлені від ядра мережі, і почалася розробка інваріантності доступу до IMS, що отримала назву IP connectivity access і передбачала застосування будь-якої

технології доступу, яка може забезпечити транспортування IP-трафіку між призначеним для користувача обладнанням і об'єктами IMS. [2]

5) Створення послуг та управління послугами

Необхідність швидкого впровадження різноманітних послуг, оскільки саме вони повинні стати основним джерелом доходів оператора в XXI столітті, вимагала негайного перегляду процесу створення послуг в IMS. Для того, щоб скоротити час впровадження послуги та забезпечити її надання в гостьовій мережі, коли користувач перебуває в роумінгу, IMS стандартизує не послуги, а можливості їх надання (service capability). Таким чином, оператор може реалізувати будь-яку послугу, що відповідає service capability, і ця послуга буде підтримуватися при переміщенні користувача до гостьової мережі, якщо ця мережа має подібні стандартизовані service capability. [2]

6) Роумінг

Функції роумінгу існували вже в мобільних мережах 2G, 3G і IMS, природно, ці функції успадкувала. Однак саме поняття «роумінг» тепер істотно розширилося і включило в себе:

- GPRS-роумінг – гостьова мережа надає RAN і SGSN, а в домашній знаходяться GGSN і IMS;
- IMS-роумінг – гостьова мережа надає IP-з'єднання і точку входу (наприклад P-CSCF), а домашня мережа забезпечує всі інші функції;
- CS -роумінг – роумінг між мережею IMS і мережею комутації каналів.

7) Безпека

Функції забезпечення безпеки є найважливішими для кожної телекомунікаційної системи, а IMS забезпечує принаймні стільки ж безпеки, скільки GPRS-мережі та мережі з комутацією каналів. IMS аутентифікує користувачів перед наданням послуги, дозволяє користувачеві вимагати конфіденційності інформації, переданої під час сеансу тощо. [2]

8) Нарахування плати

Як зазначалося вище, IMS дозволяє оператору або провайдеру послуг гнучко встановлювати тарифи для мультимедійних сесій. IMS зберігає можливість нараховувати плату за сеанс найпростішим способом – залежно від тривалості сеансу або обсягу трафіку, але також може використовувати більш складні схеми, що враховують різну призначену для користувача політику, компоненти медіа-даних, надані послуги тощо. Також потрібно, щоб дві мережі IMS, за необхідності, могли обмінюватися інформацією, необхідною для оплати сеансу зв'язку. IMS підтримує як online, так і offline оплату. [2]

1.4 Мережі мобільного зв'язку 5G

Особливість архітектури мережі 5G (5th Generation Mobile Network) полягає в тому, що традиційне поняття «архітектура мережі», заснованої на апаратних рішеннях, в мережі 5G втрачає актуальність.

Тому 5G частіше називають не мережею, а системою, або «платформою», під якою мається на увазі платформа програмна, а не апаратна. Якщо мережі 1/2/3/4G будувалися на базі апаратних рішень (обладнання), то платформа 5G будується на базі програмних рішень, зокрема, програмно-конфігурованих мереж SDN (Software Defined Network), а також віртуалізації мережевих функцій NFV (Network Function Virtualization).

Функції 5G реалізуються в віртуальних програмних функціях VNF (Virtual Network Function), які працюють в інфраструктурі NFV. Різниця між цими схожими за звучанням поняттями полягає в тому, що VNF - це функція, а NFV - це технологія. У свою чергу, NFV реалізується у фізичній інфраструктурі дата-центрів (data center, DC, центр обробки даних, ЦОД), на базі стандартного комерційного обладнання COTS (Commercial Off The Shelf). Устаткування COTS включає лише три види стандартних, відносно недорогих пристроїв - сервер, комутатор і система зберігання даних.

Архітектура опорної мережі 5G зображена на рисунку 1.6 та складається з наступних мережевих функцій (Network Function):

- функція управління доступом і мобільністю (AMF - Access and Mobility Management Function);
- функція управління сесіями (SMF - Session Management Function);
- функція передачі даних користувачів (UPF - User Plane Function);
- модуль керування даними користувачів (UDM - Unified Data Management);
- уніфікована база даних (UDR - Unified Data Repository);
- система зберігання неструктурованих даних (UDSF - Unstructured Data Storage Function);
- функція вибору мережевого шару (NSSF - Network Slice Selection Function);
- функція управління політиками (PCF - Policy Control Function);
- функція забезпечення взаємодії з зовнішніми додатками (NEF - Network Exposure Function);
- сховище втратити зв'язок із мережею (NRF - NF Repository Function);
- прикладна функція (AF - Application Function);
- функція підтримки обміну короткими текстовими повідомленнями за допомогою протоколу NAS (SMSF - SMS Function);
- функція взаємодії з не-3GPP мережею доступу (N3IWF - Non-3GPP Inter Working Function);

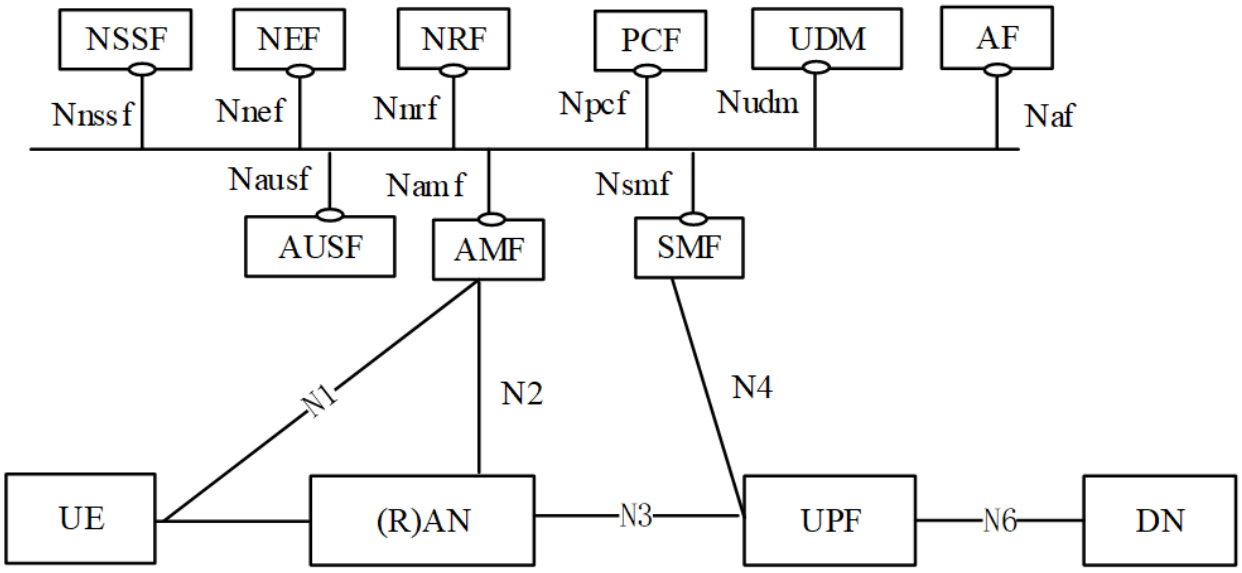

Рисунок 1.6 Архітектура опорної мережі 5G

1.4.1 Застосування концепції IMS в мережах 5G

Хоча ми дійсно вітали кінець ери передачі голосу з комутацією каналів ще на початку століття, факт полягає в тому, що навіть після титанічних зусиль між 2018 і 2019 роками кількість голосових викликів з використанням VoLTE (Voice over LTE), наприклад, в Великобританії, складає всього близько 60%. Це все не можна назвати тільки виною постачальників послуг. З 2019 року автомобільна система виклику екстрених служб (eCall) Європейського Союзу була заснована на технологіях 2G/3G, і люди не схильні купувати нові автомобілі для більш чіткого амбулаторного обслуговування. Крім того, старі телефони все ще використовуються користувачами похилого віку в сільських місцевостях всього світу і це не змінити.

Що стосується 3GPP, вони поставили рису на піску з Release 15 (5Gv1), вимагаючи кінця комутованого голосу. Але з Release 16 (5Gv2), хвиля реальності захлеснула його, і була додана можливість продовжити дозвіл на відкат до комутації каналів. У будь-якому випадку, це всього лише передісторія, тому що оператори повинні зробити все можливе, щоб підтримувати наскрізну пакетну передачу голосу в своїх блискучих нових

інфраструктурах 5G, а це вимагає поновлення їх існуючої реалізації IMS. Хоча це, очевидно, означає викорінення передачі голосу з комутацією каналів через застарілий центр комутації мобільного зв'язку (MSC - Mobile Switching Center), залишається два варіанти, в залежності від прагнення до розвитку: продовжити рух по шляху VoLTE або відразу перейти до Vo5G (Voice over 5G).

Багато з варіантів міграції підтримують більшу частину інфраструктури Evolved Packet Core (EPC), дозволяючи операторам зберегти своє існуюче ядро IMS і продовжити міграцію голосових пакетів з використанням VoLTE, якщо вони того побажають (рисунок 1.7).

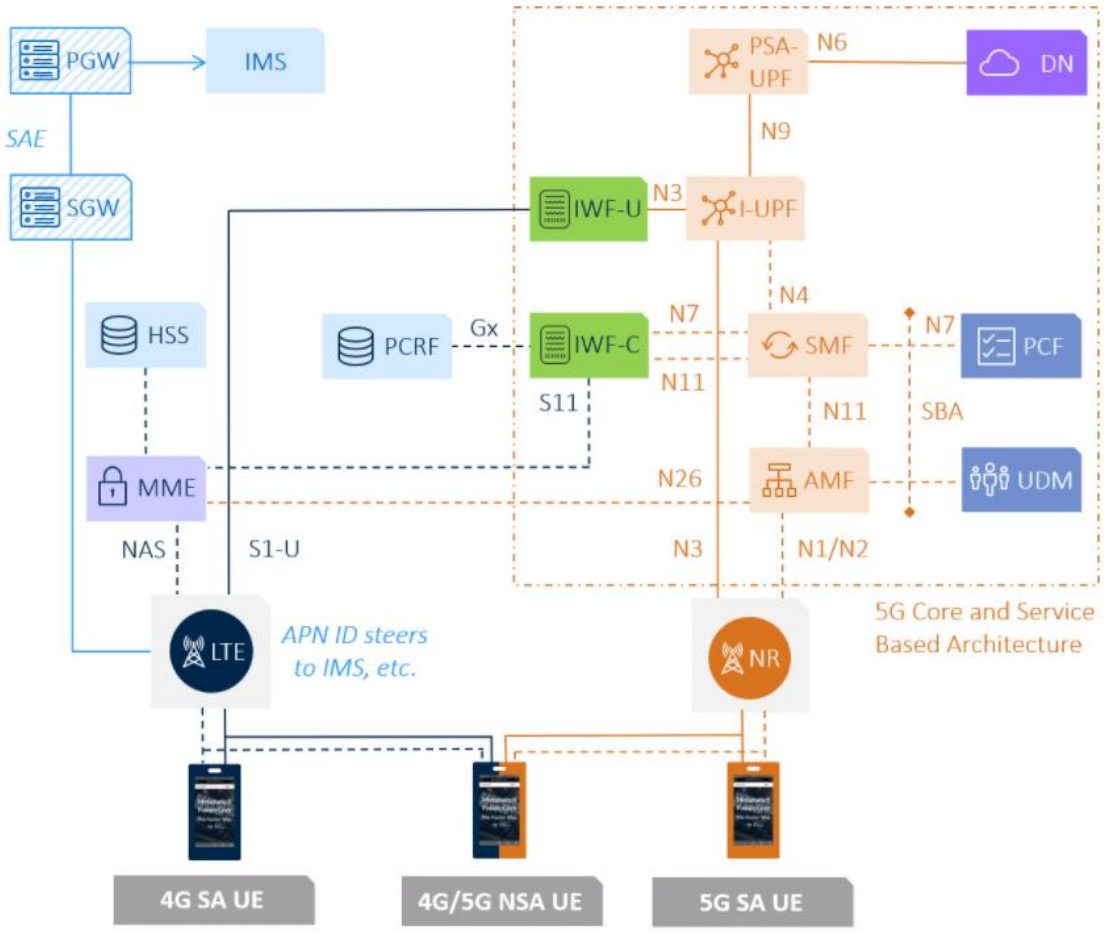

Рисунок 1.7 Підсистема IMS в 5G з підтримкою VoLTE

Сервіс оператори, звичайно, можуть підтримувати EPC, але навіщо, якщо ядро IMS з підтримкою 5G поставляється з деякими досить крутими поліпшеннями. З точки зору реалізації, хмарне ядро IMS дає можливість брати

участь в сегментації мережі, що є важливою технічною та бізнес-метою 5G. Створення високо деталізованих і оптимізованих сегментів IMS для конкретних додатків і управління трафіком за допомогою функції вибору сегмента мережі 5G (NSSF) розширює важливу концепцію ізоляції послуг на площину голосового управління.

Однак розгортання Vo5G також потребує оновлення інтерфейсів IMS, яким вже кілька десятиліть. Цьому приділяється велика увага в дослідному технічному звіті 3GPP «Дослідження вдосконаленої мультимедійної IP-підсистеми (IMS) для інтеграції 5GC», опублікованому як TR 23.794. Цей документ дуже докладно описує розширення і заміну інтерфейсів Diameter на інтерфейси на основі служб (SBI). Це дозволяє ядру 5G IMS підтримувати не тільки існуючих виробників, але більш нові мережеві функції, специфічні для 5G. Прикладом є HSS, який тепер може бути поєднаний або реалізований як частина UDM (рисунок 1.8).

1.4.2 Перспективи розвитку мереж зв'язку 5G

В останні роки еволюція мобільного зв'язку та їх інтеграція у повсякденне життя суспільства в цілому має очевидний вплив на економічний та соціальний розвиток. Мережі зв'язку 5G будуть стикатися з підключенням великої кількісті нових додатків. Управління та експлуатація мереж повинні ґрунтуватися на підключених до мережі застосунках. Мережі, що базуються на додатках, складаються з взаємопов’язаних пристроїв, різних модулів, машин, датчиків і виконавчих механізмів та безлічі клієнтів, підключених до Інтернету, які генерують великі потоки трафіку. Дизайн мережевої технології 5G повинен відповідати потребі у розробці нових мережевих моделей, які можуть бути відкритими, більш гнучкими та масштабованими. Майбутні комунікаційні мережі повинні спростити процес створення мережі відповідно до вимог конкретних послуг, що надаються (рисунок 1.9).

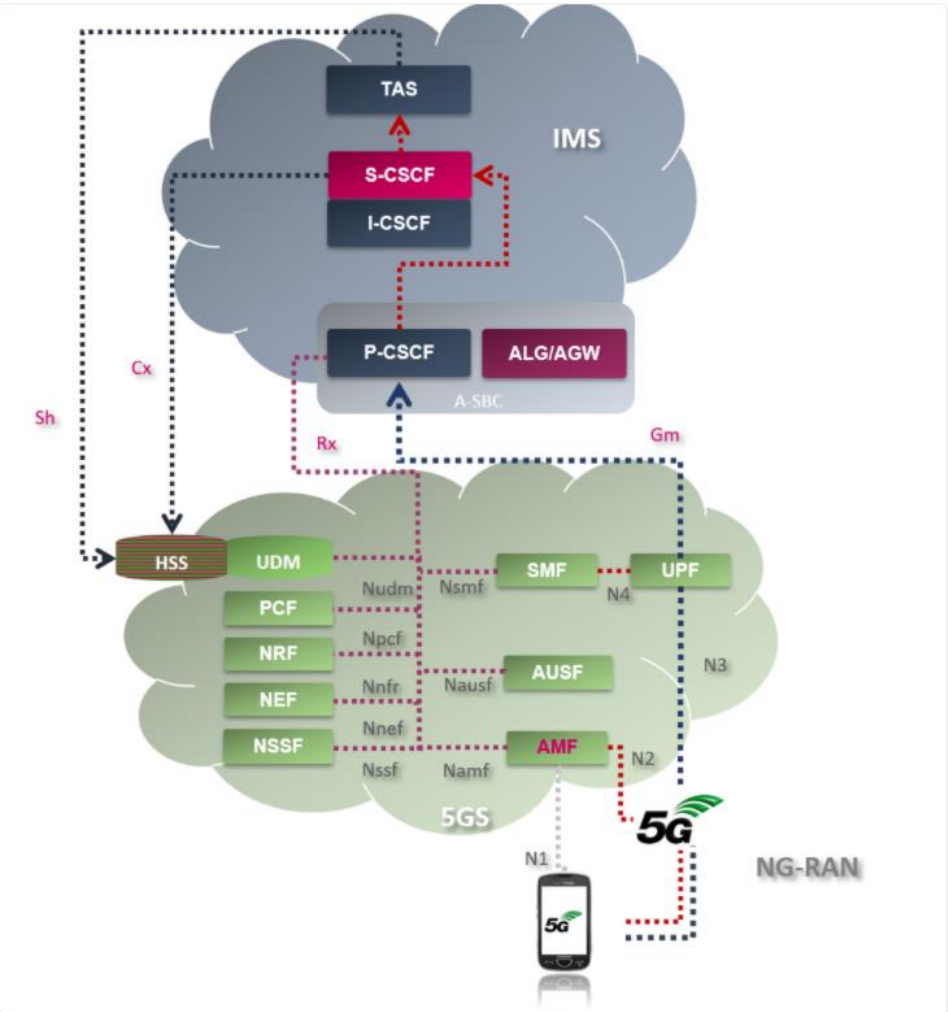

Рисунок 1.8 Підсистема IMS в 5G з підтримкою Vo5G

Мережі зв'язку 5G забезпечуватимуть конвергентний зв'язок в мережах з різними технологіями і відкриту систему зв'язку для взаємодії з супутниковими системами, стільниковими мережами, хмарними системами, інформаційними центрами, домашніми шлюзами і багатьма іншими відкритими мережами і пристроями. Крім того, мережі зв'язку 5G будуть автономними і програмованими. Відповідно, безпека, відмово стійкість, надійність і цілісність даних будуть пріоритетними при проектуванні майбутніх мереж. Крім цього, мережі зв'язку 5G повинні забезпечити мобільність користувачів для можливості з’єднання в будь-якому місці, в будь-який час та в будь-яких поєднаннях. [4]

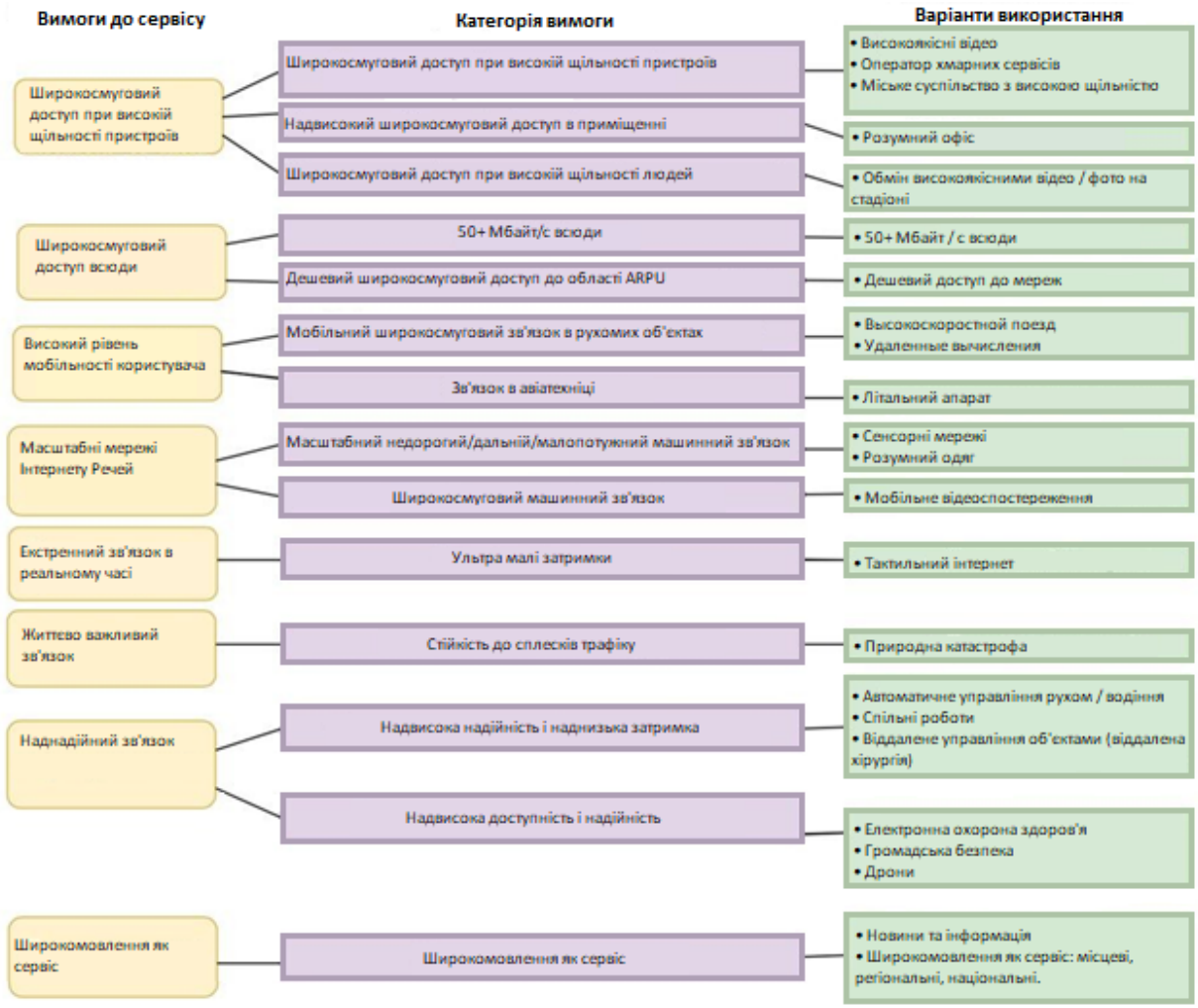

Рисунок 1.9 Концептуальний сценарій реалізації мереж 5G

1.4.3 Необхідність програмування мереж зв'язку 5G

Для досягнення необхідної гнучкості спочатку необхідно, щоб існуючі ресурси в інфраструктурі могли бути адаптовані і змінені динамічно. Крім того, потрібні методи і механізми для розробки додатків і сервісів поверх гнучкої інфраструктури в різних технологічних областях. Для досягнення цих вимог потрібна архітектура управління з високим рівнем програмовості. Зокрема, архітектура управління не повинна бути прив'язана до конкретного варіанту використання або сценарію, і повинна дозволити мережевому операторові програмувати настроювані алгоритми на рівні управління для оптимізації

мережі радіодоступу RAN (Radio Access Network), транспортної мережі та мережевих ресурсів хмарних платформ. Додатково необхідні наступні функції:

1) модульність. Архітектура управління повинна відповідати модульній архітектурі з чітко визначеними функціями управління та інтерфейсами. Інтерфейси та архітектурні будівельні блоки також повинні підтримувати стикування рекурсивно, щоб дозволити реалізацію системи, адаптованої до конкретних сценаріїв;
2) віртуалізація. Архітектура повинна мати можливість відокремлювати фізичні та віртуальні ресурси інфраструктури на окремі групи (або фрагменти) і розподіляти їх між різними клієнтами. Тут клієнтами можуть бути контролери вищого рівня з функцією обслуговування різних додатків. Відокремлені фрагменти мережі слід ізольовати один від одного через проблеми безпеки і продуктивності (наприклад, для запобігання негативного впливу перевантаженого фрагмента мережі на інші фрагменти);
3) масштабованість. Управління ресурсами в кожному з доменів є складним завданням, оскільки ми, як правило, маємо справу з великою кількістю елементів мережі, а також з параметрами та процедурами управління. Таким чином, спільний контроль над доменами може легко стати нерозв'язним, чого слід уникати при правильному проектуванні архітектури управління. Також необхідні відповідні методи абстракції, щоб обмежити складність в вищих шарах та зробити загальну проблему оптимізації керованою. Для задоволення цих вимог при розробці загальної архітектури управління застосовуються принципи організації мережі, засновані на технології програмно-конфігурованих мереж (ПКМ) і віртуалізації мережевих функцій.

1.4.4 Віртуалізація функцій в мережах зв'язку 5G

Організація мереж з програмованими параметрами і віртуалізацією функцій мережі представляють собою майбутнє, в якому віртуальна інфраструктура і послуги забезпечують безпрецедентну гнучкість, інтелектуалізацію і відкритість.

Протягом останніх п'яти років технології програмно-конфігурованих мереж і віртуалізації мережевих функцій удосконалювалися завдяки унікальній взаємодії організацій по стандартизації з спільнотами розробників програмного забезпечення з відкритим початковим кодом, які разом змінюють методи сприймання нової технології.

Інноваційні галузеві групи, такі як Робоча група ISG ETSI (Industry Specification Group of European Telecommunications Standards Institute) по NFV і організація ONF (Open Networking Foundation) створили еталонні архітектури, обґрунтували сценарії використання і змінили вимоги до складових елементів з відкритим початковим кодом, які є невід'ємною частиною NFV і SDN.

Технології програмно-конфігурованих мереж і віртуалізації мережевих функцій, стали найважливішим ключем для сучасних мереж, що сприяв розробці широкого кола додатків, в тому числі додатків для операторів мобільного зв'язку, Інтернету речей і т.п..

Для того, щоб забезпечити можливість використання такого широкого кола застосунків для кінцевих користувачів, модель управління і контролю SDN/NFV повинна стати набагато більш масштабованою, інтелектуальною, гнучкою і відкритою, ніж будь-коли раніше (Рисунок 1.10).

З цих причин найбільш прогресивні і ініціативні оператори і постачальники мережевих послуг взяли на себе завдання зміни життєвого циклу процесу надання послуг. Для цього необхідна безпрецедентна взаємодія організацій по стандартизації, операторів, галузевих організацій та спільноти розробників програмного забезпечення з відкритим вихідним кодом.

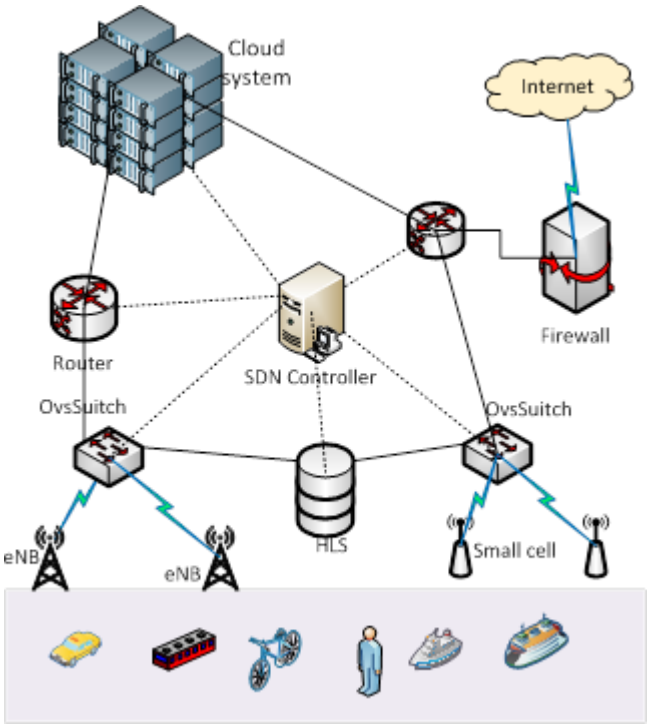

Рисунок 1.10 Загальна архітектура SDN в мобільних мережах

1.5 Концепція сегментації в мережах зв'язку

Кількість пристроїв та користувачів мобільних мереж постійно зростає, з'являються все нові і нові мережеві послуги, вимоги абонентів до швидкості доступу в інтернет загалом та мобільного доступу до інтернету, зокрема, стають жорсткішими і постійно збільшуються, абоненти стають більш вимогливими до якості обслуговування в цілому. Розробники програмного і апаратного забезпечення мереж і оператори зв'язку, бажаючи відповісти на нові завдання, трансформують архітектуру мереж і норми, стандарти, регламенти взаємодії. Таким чином, нещодавно з'явилися мережі п'ятого покоління (5G), які є серйозним продовженням розвитку мереж четвертого покоління (LTE, 4G).

У мережах зв'язку 5G найважливішою концепцією є мережева сегментація. Взагалі технологія слайсингу визначається як концепція паралельного розгортання кількох логічних, автономних і незалежних мереж на спільній інфраструктурній платформі. В принципі, сегментація мережі є

специфічним способом віртуалізації мережевих інфраструктурних ресурсів; це інструмент для спільного використання мережевих ресурсів і надання настроюваних мережевих архітектур для різних цілей, що використовують одну і ту ж базову фізичну інфраструктуру [9]. Кожна така віртуальна мережа – це сегмент мережі, який має свій тип трафіку і який може використовувати власну технологію передачі даних. Гнучкість сегментації мережі дозволяє задовольнити найрізноманітніші і навіть суперечливі вимоги абонентів. Завдяки концепції слайсингу оператори зв'язку можуть ізолювати мережеві ресурси «за вимогою» на основі спільної фізичної інфраструктури.

Таким чином, у сегментованій мережі кожен сегмент є окремою логічною мережею, налаштовану під певні сервіси. Мережевий сегмент відноситься до керованих розділів фізичних і/або віртуальних мережевих ресурсів; до фізичних, віртуальних та сервісних функцій мережі; які можуть виступати як незалежний екземпляр мережі і/або як мережева хмара. До мережевих ресурсів належать комунікаційні ресурси (канали зв'язку, телекомунікаційне обладнання тощо), обчислювальні ресурси і ресурси зберігання.

Згідно [10], реалізації мережевої сегментації в мережах зв'язку 5G - переважно для мережевих адміністраторів і дизайнерів, завдяки в основному наступним наборам функцій:

1) ізоляція мережевих сегментів: повна ізоляція сегмента мережі надає можливість керувати паралельними незалежними мережевими сегментами. В результаті чого збої мережі, перевантаження або інші загрози безпеки в межах одного сегменту не впливають на роботу інших в масштабах мережі. Більше того, кожен сегмент мережі повинен мати незалежні функції захисту, які зупиняють несанкціоновані спроби доступу на читання або запис інформації про конфігурацію, управління або обліку;
2) гнучка віртуалізація мережевих функцій: на відміну від традиційних мобільних мереж, в яких всі сервіси складаються з однакових функцій, за

допомогою технології сегментації кожна служба може залежати від інших функцій;

3) спрощені сервісні ланцюжки в домені віртуалізації: технологія NFV дозволяє використовувати мережеві функції незалежно від їх фізичного розташування. Однак оптимальне розміщення мережевих функцій покращує можливості мережі;
4) прозоре управління мережевими сегментами: завдяки рівню абстрагування фізичних мережевих ресурсів, технологія сегментації в мережах зв'язку дозволяє встановлювати мережеві сегменти для мереж різних доменів.

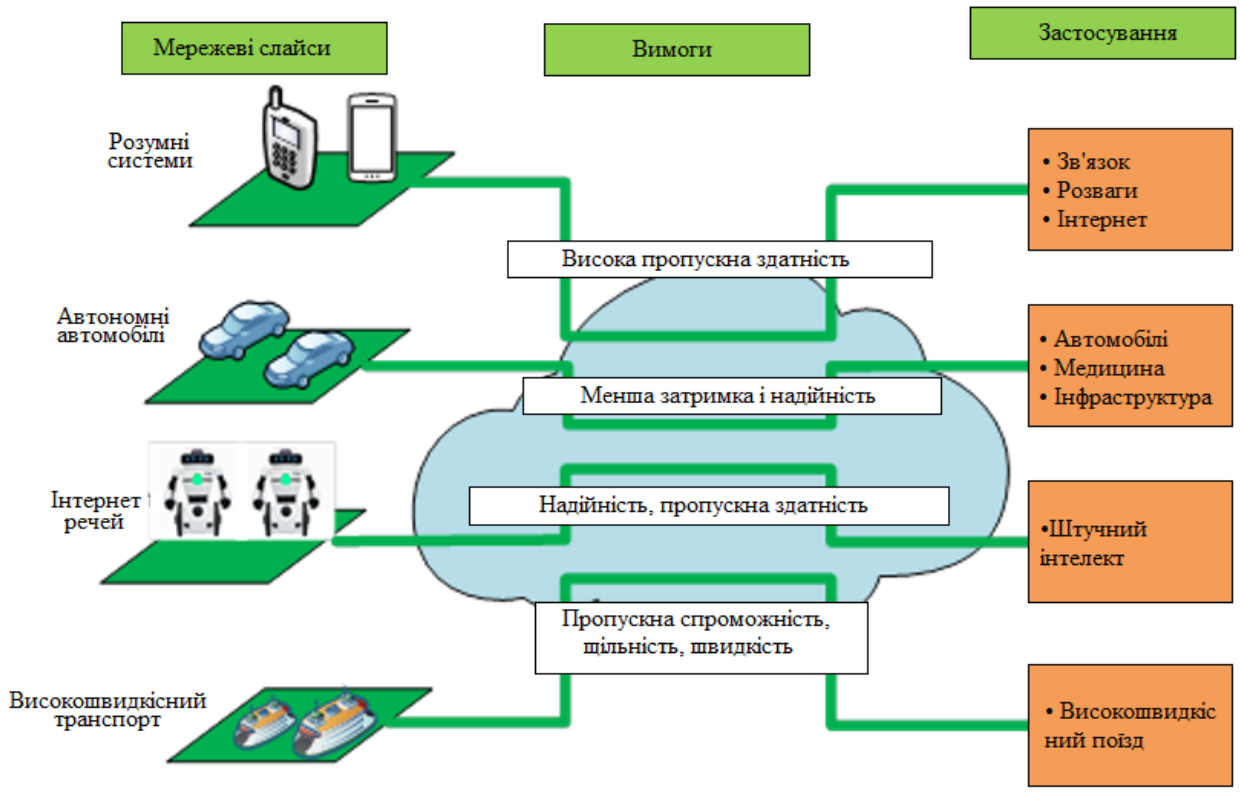

Рисунок 1.11 Віртуальні функції в мережах зв'язку 5G і відповідні їм вимоги для мережевих сегментів

Один сегмент мережі спеціально орієнтований на забезпеченні низької затримки та високої надійності (наприклад, автономні транспортні засоби). В свою чергу інший мережевий сегмент призначений для пристроїв з низьким

енергоспоживанням (наприклад, датчики), а також акумуляторів малої ємності, а інший для сервісів, що вимагають надвисокі швидкості.

1.6 Висновки до розділу 1

У першому розділі розглядаються передумови появи мереж NGN та їх еволюція до мультисервісних IP мереж. Таким чином, пункти цього розділу містять інформацію про те, як ця архітектура виникла в результаті еволюційного процесу мереж NGN, перейнявши від них ідею функціонування на основі технології комутації пакетів та забезпечення конвергентної роботи послуг, що в ній надаються. Досліджена структура (IMS) та її роль як у сучасних інфокомунікаційних мережах, так і необхідність подальшого застосування у майбутньому.

Також, в рамках даного розділу наводиться інформація про будову майбутніх мобільних мереж 5G та їх безпосередній зв’язок з підсистемою IMS. Тут з’ясовується, що з конвергенція цих технологій дозволяє забезпечити надання широкого спектру нових послуг усім користувачам мережі, незалежно від їх розташування, або ж кінцевого обладнання яке вони використовують, при забезпечуючи при цьому максимально високу якість обслуговування для кожного з абонентів. Разом з мережами 5G вводиться поняття програмно-конфігурованих мереж (SDN) і віртуалізації мережевих функцій (NFV). Дані технології дозволяють реалізувати основну бізнес ідею мереж 5 покоління – сегментування. Яка, в свою чергу, забезпечує гнучкість мережі, розбиваючи одну фізичну мережу на кілька шарів, кожен з яких має власні налаштування, адаптовані під певну послугу. Таким чином забезпечується ефективність і гнучкість майбутніх сервісів.

В результаті роботи над розділом, виясняється, що гнучке управління мережею за допомогою технології SDN та віртуалізовані мережеві функції технології NFV допоможе розмити границі між вендорами, розв’язати проблему росту пакетного трафіку та напряму допоможе операторам

раціоналізувати витрати на обслуговування і побудову мережі, що, в свою чергу, автоматично збільшить дохід. Віртуалізація і програмованість їхніх мереж дозволить задовольняти потреби кожного з користувачів, шляхом гнучкого управління мережею та надання абонентам тих послуг, які їм потрібні, не переплачуючи.

2 АНАЛІЗ КОНЦЕПЦІЇ ПРОГРАМНО-КОНФІГУРОВАНИХ МЕРЕЖ ТА ПРОТОКОЛУ OPENFLOW

2.1 Розвиток програмно-конфігурованих мереж

Незважаючи на те, що технологія програмно-конфігурованих мереж на перший погляд виникла недавно, насправді вона має досить тривалу передісторію. SDN – це завершальна фаза тривалого періоду зусиль, спрямованих на те, щоб зробити мережі програмованими. Телекомунікаційні мережі завжди відрізнялися складністю структури і управління. Для їх побудови використовується безліч різноманітного обладнання: комутатори і маршрутизатори, а також брандмауери, транслятори мережевих адрес, балансувальники навантаження, тощо. Зазвичай мережеві адміністратори налаштовують кожен пристрій окремо, дотримуючись мережевої топології і правил маршрутизації. Для цього використовуються інтерфейси для конфігурації – свої для кожного пристрою і для кожного виробника, а іноді і для кожного пристрою всередині лінійки одного виробника.

Програмно-конфігуровані мережі змінюють підхід до проектування та адміністрування мереж. По-перше, SDN відокремлює площину управління мережею (Control plane), яка займається маршрутизацією трафіку, від площини передачі даних (Data plane), яка передає трафік згідно з, отриманими від площини управління, правилами. По-друге, SDN «консолідує» площину управління, при цьому один комплекс керуючих програм на сервері керує багатьма пристроями на площині даних. Для цього використовується стандартизований інтерфейс прикладного програмування API. [5] Це, іншими словами, такий інтерфейс, як OpenFlow, детально ми будемо розглядати його далі. Виходячи з цього, для побудови мережі SDN, на елементах мережі, перш за все, комутаторах і маршрутизаторах, повинна бути реалізована підтримка OpenFlow. При цьому на кожному з них є таблиця, або таблиці, правил маршрутизації. Кожне правило визначає, як маршрутизувати пакети певної сесії або потоку трафіку. Виходячи з правил, встановлених керуючою прикладною

програмою, кожен комутатор OpenFlow може працювати як комутатор, маршрутизатор, брандмауер, транслятор мережевих адрес, тощо.

Незважаючи на те, що концепція SDN стала популярною протягом останніх років, сама ідея досить стара, і еволюціонує вже більше 2-ух десятиліть. Її сліди можна простежити навіть у розвитку ранніх телефонних мереж на базі комутації каналів, коли управління мережею було відокремленою від мережі канальної комутації. І це було зроблено рівно з тією ж метою, що і в SDN - щоб спростити управління і введення нових послуг. Концепція так званих «гнучких комутаторів» Softswitch для телекомунікаційних мереж на базі комутації пакетів також дуже близька до SDN за функціями і реалізацією.

У міру збільшення зацікавленості до функціональності ПКМ, велика увага при побудові архітектури змістилася з гнучкої передачі пакетів на динамічну віртуалізацію ресурсів та оркестрацію сервісів. В результаті чого отримана архітектура може бути застосована до всіх видів додатків в мережах підприємства, операторів, центрів обробки даних і мережах кампусів, від кінцевого споживача до власника апаратного забезпечення, як для абсолютно нових, так і для існуючих мережевих реалізацій, що адаптує ПКМ до різних сфер застосування.

2.2 Архітектура і принципи побудови програмно-конфігурованих мереж

Традиційні комп'ютерні мережі складаються з взаємопов'язаних мережевих пристроїв, такі як комутатори і маршрутизатори (рисунок 2.1). Кожен пристрій має механізм передачі на транспортному рівні та систему управління, яка включає в себе операційну систему і додатки. У цій моделі мережеві пристрої мають закриту архітектуру і немає можливості додати нові функції. Прив'язка мережевих пристроїв до вибраного мережевого виробника не гарантує підтримку майбутніх програм і сервісів.

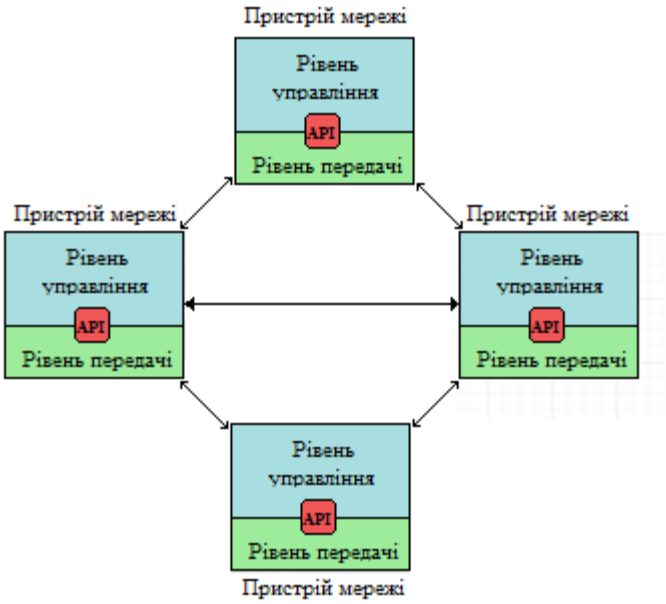

Рисунок 2.1 Структура традиційної мережі передачі даних

Архітектура програмно-конфігурованих мереж визначена в ONF TR-502 [6], та базується на наступних трьох принципах:

1) поділ функцій управління і передачі трафіку.
2) централізоване управління – це означає, що зовнішнє управління (клієнт/додаток) виглядає як єдине ціле, що гарантує ефективне використання ресурсів мережі порівняно з перспективою обмеженого управління;
3) програмованість сервісів мережі. Інтерфейси між об’єктами SDN розкривають абстракції ресурсів і стан мережі, а також забезпечують обмін інформацією та управління мережею. Додатки можуть встановлювати вимоги і просити зміни в своїх мережевих служб, а також програмно реагувати на стан мережі.

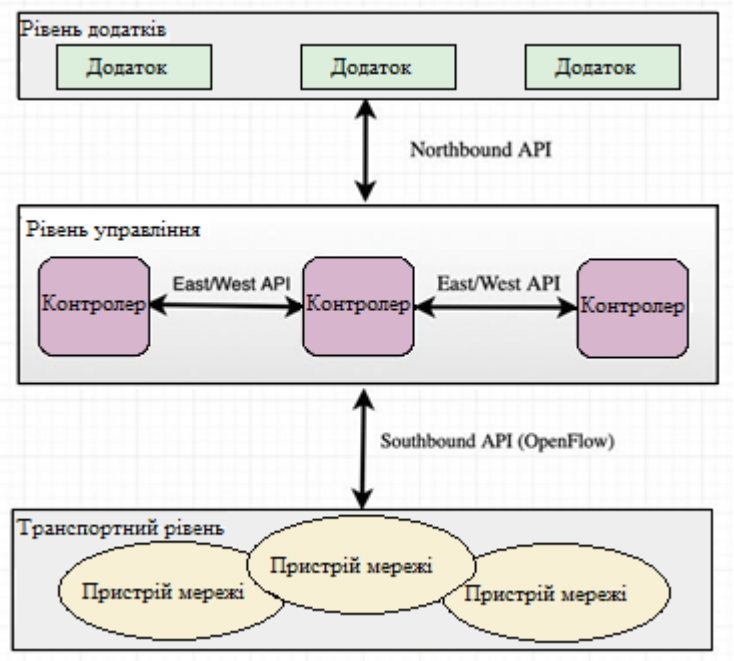

Рисунок 2.2 Структура мережі SDN

Як показано на рисунку 2.2, архітектура ПКМ містить шість основних компонентів.

1) Рівень додатків. На верхньому рівні, який називається рівнем додатків або прикладним рівнем, реалізуються різні служби, такі як системи виявлення та запобігання вторгнень (IDS/IPS - Intrusion Detection Systems/Intrusion Prevention Systems), системи забезпечення якості обслуговування (QoS), контроль доступу, проксі-сервер і інші, які в свою чергу також покладаються на ресурси мережі.
2) Рівень управління. Рівень управління абстрагує топологію мережі від рівня додатків за допомогою northbound API (північного API-інтерфейсу). На цьому рівні основним елементом є контролер SDN, який відповідає за координацію одного або декількох мережевих пристроїв на рівні передачі даних. Контролер контролює обробку даних для управління елементами мережі, відстежує їх стан і здійснює збір інформації про мережу через southbound API (південний API-інтерфейс). Протокол OpenFlow є прикладом реалізації такого API. На цьому рівні існують також такі елементи, як eastbound і westbound API (східний і західний API-

інтерфейси), які забезпечують зв'язок з іншими контролерами мережі та дозволяють їм обмінюватися інформацією, пов'язаною з обробкою трафіку на транспортному рівні

3) Транспортний рівень. Нижній рівень, відомий як транспортний рівень або рівень передачі даних, забезпечує обробку і пересилку пакетів на підставі отриманих від рівня управління інструкцій. На цьому рівні знаходяться пристрої мережі, якими керує контролер.
4) Northbound interfaces. Північні API-інтерфейси представляють собою мережеві інтерфейси між SDN-додатками та SDN-контролерами. Саме ці інтерфейси найбільш важливі API-інтерфейси архітектури ПКМ, так як вони дозволяють додаткам використовувати сервіси мережі та динамічно її налаштувати. Мережеві сервіси, які можуть бути розгорнуті і оптимізовані за допомогою північних API-інтерфейсів, включають сервіси безпеки, балансування навантаження, управління трафіком, якістю обслуговування, тощо.
5) Southbound interfaces. Південні API-інтерфейси забезпечують ефективне управління мережею і дозволяють SDN-контролеру динамічно організовувати ресурси мережі відповідно до вимог та потреб в реальному часі. Південні API-інтерфейси на відміну від північних взаємодіють з пристроями мережі (комутатор, маршрутизатор). З різних протоколів південного API-інтерфейсів найбільш використовуваний - OpenFlow.
6) East/West interfaces. Східні/Західні API-інтерфейси забезпечують зв'язок між SDN-контролерами і дозволяють їм обмінюватися інформацією, яка стосується обробки трафіку на рівні передачі даних.

2.2.1 Контролер програмно-конфігурованої мережі

Контролер є найбільш важливим компонентом архітектури SDN, вузлом, який централізує функції мережі і відстежує її глобальний стан. Контролер SDN

може бути представлений як апаратно-програмне рішення, так і в якості програмної реалізації. На даний момент, контролери на ринку частіше представлені в програмній реалізації. Видаляючи рівень управління з мережевого обладнання та запускаючи його як програмне забезпечення, контролер спрощує автоматичне керування мережею, а також інтеграцію і адміністрування бізнес-додатків. Перші варіанти архітектури ПКМ представляють план управління, що складається з єдиного централізованого контролера. Пізніше були запропоновані розподілені архітектури з використанням декількох контролерів для підвищення продуктивності і масштабованості мережі.

Серед основних характеристик контролера програмно-конфігурованих мереж можна виділити:

1) продуктивність - число потоків, оброблюваних контролером за одиницю часу [потік/с];
2) час обробки - кількість часу, що витрачається контролером на обробку запиту від комутатора [с];
3) надійність - число відмов при заданому профілі навантаження;
4) ресурсомісткість - утилізація контролером оперативної пам'яті фізичного сервера і навантаження на ядра процесора;
5) масштабованість - підтримка контролером багатопотоковості.

SDN-контролер зазвичай містить набір модулів, які при підключення виконують різноманітні мережеві завдання. Основні модулі включають в себе:

- модуль виявлення каналів;
- модуль топології;
- модуль пам'яті;
- модуль вироблення стратегії;
- модуль таблиці потоків;
- модуль управління даними.

Також існують модулі які виконують складніші функції.

За надання послуги маршрутизації відповідають два модулі: модуль топології і модуль виявлення каналів. Модуль виявлення каналів відповідає за виявлення і підтримку стану фізичних з'єднань в мережі. Процедура запускається контролером, коли будь-який невідомий тип трафіку потрапляє в домен OpenFlow. Таким чином, інформація, зібрана модулем виявлення каналів, використовується для створення бази даних сусідів на контролері. З цієї бази даних модуль топології створює, а згодом, підтримує оновлення інформації про топологію і обчислює маршрути в мережі.

На сьогоднішній день існує ряд контролерів, більшість з яких мають відкритий вихідний код і підтримують протокол OpenFlow. Ці контролери відрізняються мовами програмування, підтримуваною версією OpenFlow, методами, що використовуються в якості багато потоковості і продуктивністю в якості бітрейта.

2.2.2 Протокол OpenFlow

Протокол OpenFlow є відкритим стандартом. Він був запропонований ONF для стандартизації взаємодії контролера з мережевими пристроями в архітектурі ПКМ. Слідуючи еволюційному шляху розвитку телекомунікаційних мереж, протокол OpenFlow, на даному етапі розробки і практичної реалізації концепції ПКМ, реалізується у вигляді програмного модуля в апаратних реалізації Ethernet комутаторів, маршрутизаторів і бездротових вузлів доступу, як продовження можливості їх застосування в якості пристроїв ПКМ [7].

В даний час стандарт OpenFlow де-факто приймається більшістю виробниками мережевого обладнання. На телекомунікаційному ринку на даний момент доступні комутатори з підтримкою OpenFlow від таких виробників як Cisco, Juniper, Brocade, Huawei, Zelax та інші.

Повідомлення протоколу OpenFlow прийнято ділити на три типи:

1. Повідомлення контролер-комутатор – ініціалізуються на контролері, служать для управління комутатором і контролем за подіями, що

відбуваються на ньому. До даного типу повідомлень відносяться наступні повідомлення:

- *Features*: це повідомлення використовується для запиту контролером можливостей комутатора; комутатор в свою чергу відповідає на такий запит повідомленням features, в якому вказує свої можливості. Цей процес відбувається під час відкриття каналу OpenFlow;
- *Configuration*: цим повідомленням контролер робить запит і встановлює параметри налаштувань комутатора;
- *Modify-State*: ці повідомлення надсилаються ОС мережі для управління станом комутаторів. Задача повідомлень: додавання, видалення правил і змінення OpenFlow таблиць, настройка портів комутатора;
- *Read-State*: дані повідомлення відповідають за збір статистики комутаторів;
- *Packet-out*: повідомлення Packet-out використовуються контролером для відправки пакетів з певного порту комутатора і пересилку пакетів, отриманих за допомогою повідомлення Packet-in, які містять весь пакет або ідентифікатор ID буфера, що відноситься до пакета, завантаженого в комутатор. Повідомлення повинно містити список дій, які застосовуються у зазначеному порядку: якщо список дій порожній, то пакет скидається;
- *Barrier*: повідомлення Barrier забезпечують встановлення залежностей між повідомленнями або повідомлення про завершення операції. Використовуються при необхідності обробки повідомлень в певному порядку;
- *Role-Request*: даний запит слугує для зміни пріоритету контролера на комутаторі (для підвищення ролі з Slave до Master);

- *Asynchronous-Configuration*: за допомогою даного повідомлення контролер встановлює фільтр на асинхронні повідомлення від комутаторів.

2. Другим типом повідомлень є асинхронні повідомлення, які ініціалізуються OpenFlow-комутаторами, призначені для сповіщення контролера про події на мережі, наприклад, збої, помилки, зміни стану. До даного типу повідомлень відносяться такі повідомлення:
   - *Packet-in*: дане повідомлення ініціюється комутатором і надсилається контролеру в разі, якщо отриманий пакет не має необхідного правила в таблиці комутації. Для всіх пакетів, які пересилаються в віртуальний порт, повідомлення Packet-in відправляється на контролер;
   - *Flow-Removed*: повідомлення для видалення правил, які не використовуються і неактивні;
   - *Port-status*: генеруються комутатором на контролер в разі зміни налаштувань порту;
   - *Error*: цим повідомленням контролер сповіщає про помилки або збої.
3. Третім типом повідомлень є симетричні повідомлення, які розсилаються як комутаторами, так і контролером. До даного типу повідомлень відносяться наступні повідомлення:
   - *Hello*: повідомлення, обмін якими відбувається між комутатором і контролером при встановленні з'єднання;
   - *Echo*: повідомлення типу запит/відповідь можуть ініціюватися і контролером, і комутатором, при умові, що обов'язково буде отримана відповідь. Також можуть служити для виміру затримок або пропускної здатності з'єднання контролер-комутатор, а також перевірки ефективності з'єднання;

- *Experimenter*: повідомлення Experimenter призначені для забезпечення додаткової функціональності, при проведенні експериментів в просторі типів повідомлень OpenFlow.

2.2.3 Протоколи NB-API

До середини 2013 р. ONF цілеспрямовано не розробляв і не накладав вимог на NB-API (Northbound Application programming Interface, північний програмний інтерфейс), що реалізується контролерами для організації взаємодії з мережевими програмами. Рішення, призначені для реалізації функцій контролерів, пропонують власні API, конкуруючі в ринкових умовах.

Сформована тенденція в практичних рішеннях SDN пропонує два типи реалізації NB-API інтерфейсу:

1) REST (або RESTful - Representational State Transfer) - архітектурне рішення взаємодії компонент розподіленого додатка в мережі, на основі моделі «клієнт - сервер», в основі зазвичай використовується протокол HTTP і формати передачі даних - * .xml, * .json;

2) програмний інтерфейс взаємодії програмного забезпечення (ПЗ) (наприклад, Java API платформи Java EE або Java OSGI API), який безпосередньо залежить від реалізації контролера SDN.

Зазвичай, як в першому (REST API), так і в другому випадку, існує документація на наданий програмний інтерфейс, а також можливі відповіді системи (в даному випадку контролера) на відповідні запити. Тим самим SDN контролер з точки зору адміністратора або стороннього розробника додатків під контролер, виглядає як «чорний ящик» з відомим набором функцій і реакцій на них.

2.2.4 Порти OpenFlow

Пакети OpenFlow приймаються вхідним портом і обробляються конвеєром, який може спрямувати їх на вихідний порт, використовуючи вихідні дії, які визначають, як пакет повинен повертатися назад в мережу. Порти OpenFlow є мережевими інтерфейсами для передачі пакетів між обробкою OpenFlow і рештою мережі.

Комутатор OpenFlow повинен підтримувати три типи портів: фізичні, логічні і зарезервовані;

1) фізичні порти - порти, що відповідають апаратним інтерфейсам комутатора;
2) логічні порти - порти, які не відповідають безпосередньо апаратним інтерфейсам комутатора. Обробка, виконувана логічним портом, залежить від реалізації і повинна бути прозорою для обробки OpenFlow;
3) зарезервовані порти - порти, що визначають загальні дії пересилки, такі як відправка контролеру, розсилка чи пересилання за допомогою методів, відмінних від OpenFlow. Включають в себе обов'язкові порти: ALL, CONTROLLER, TABLE, IN_PORT, ANY, UNSET і опціональні: LOCAL, NORMAL, FLOOD.

2.2.5 Канал OpenFlow

Канал OpenFlow використовується для обміну повідомленнями між комутатором OpenFlow і контролером OpenFlow. Контролер може керувати кількома каналами, кожен з яких призначений для різних комутаторів. У свою чергу, комутатор також може мати канали для кількох контролерів.

Канал OpenFlow зазвичай створюється як одне мережеве з'єднання між комутатором і контролером, використовуючи протокол TLS або звичайний TCP, що гарантує підтримку з'єднань OpenFlow в широкому діапазоні мереж і умов. Однак, канал може складатися з декількох допоміжних мережевих

з'єднань на основі протоколів TLS (Transport Layer Security), TCP (Transmission Control Protocol), DTLS (Datagram Transport Layer Security) і UDP (User Datagram Protocol).

Кожне з'єднання підтримується окремо, якщо з'єднання з певним контролером або комутатором перервано, це не призводить до завершення з'єднання з іншими контролерами або комутаторами. Однак, якщо первинне з'єднання завершено або розірвано, усі відповідні допоміжні з'єднання також завершуються.

При перериванні з'єднання через мережеві умови або таймаута, комутатор або контролер, в залежності від того, хто був ініціатором з'єднання, намагається підключитися до іншої сторони до тих пір, поки не буде встановлено нове з'єднання або повністю не буде видалено його URI (Uniform Resource Identifier) з'єднання з конфігурації. Якщо комутатор втрачає контакт з одним або декількома контролерами, він також надсилає повідомлення Controller-Status (статус контролера) іншим підключеним контролерам і, коли з'єднання відновлюється, надсилає їм оновлене повідомлення про стан контролера.

У випадку, якщо комутатор втрачає контакт з усіма контролерами, він повинен негайно перейти в «fail secure mode» (пакети і повідомлення які призначені для контролера видаляються), або «fail standalone mode» (обробляє всі пакети, використовуючи зарезервований порт, діє як звичайний Ethernet-комутатор), залежно від реалізації та конфігурації комутатора.

2.3 Віртуалізація функцій мережі

Згідно з рекомендацією Міжнародного Союзу Електрозв'язку, віртуалізація мережі - технологія, яка дозволяє створювати логічно ізольовані ділянки мережі в рамках спільно використовуваних фізичних мереж таким чином, що в цій спільно використовуваній мережі одночасно можуть співіснувати багато віртуальних підмереж (LINP - Logically Isolated Network Partition, логічно ізольовані ділянки мережі). При цьому, говорячи про

віртуалізацію мережі, варто відзначити таке поняття, як віртуальний ресурс (VR - Virtual resource) - це абстракція фізичного або логічного ресурсу, яка може мати характеристики, що відрізняються від характеристики цього фізичного або логічного ресурсу, і її функціональні можливості можуть бути не пов'язані з функціональними можливостями самого фізичного або логічного ресурсу.

Таким чином, NFV - технологія віртуалізації фізичних елементів (фізичних ресурсів) телекомунікаційної мережі шляхом виконання мережевих функцій програмними модулями, що працюють на стандартних серверах і віртуальних машинах (VR - Virtual resource) в них. При цьому дані програмні модулі можуть взаємодіяти між собою для надання послуг зв'язку, чим раніше займалися апаратні рішення.

Різні типи мережевого обладнання можуть бути розташовані на стандартних промислових серверах великих обчислювальних потужностей, комутаторах і систем зберігання, які можуть бути розташовані в центрах обробки даних, мережевих вузлах і в приміщеннях кінцевих користувачів. «Віртуалізація» включає в себе реалізацію функцій мережі в програмному забезпеченні, яке може працювати в масштабі промислового сервера і за вимогою бути переміщеними в різні місця мережі, без необхідності установки нового обладнання, рисунок 2.3 ілюструє архітектуру високого рівня NFV за визначенням ETSI [8].

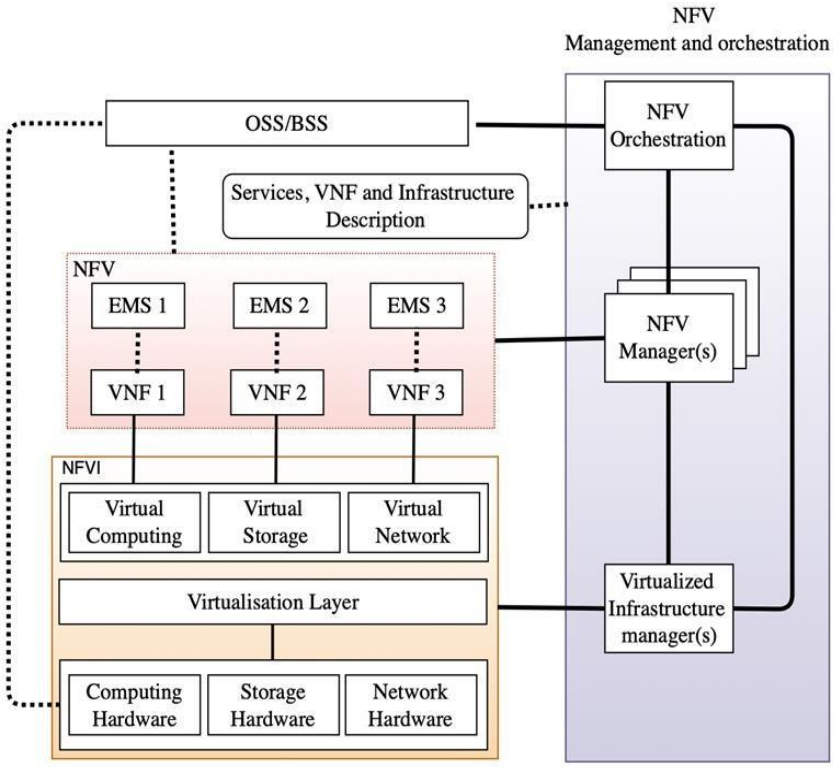

Рисунок 2.3 Архітектура високого рівня NFV

Архітектура високого рівня NFV складається з наступних трьох основних функціональних блоків:

1) Virtualized Network Function (VNF) – функції віртуалізованої мережі VNF є програмною реалізацією мережевих функцій, які зазвичай реалізовані у вигляді апаратного рішення. Наприклад, такі функції, як: Mobility Management Entity (MME), Serving Gateway (SGW), Packet Data Network Gateway (PGW), Dynamic Host Configuration Protocol server (DHCP-сервер), firewalls, DPI і інші, можуть бути реалізовані в якості програмної реалізації, які працюють на сервері(-ах). Варто також зауважити, що функції віртуалізованої мережі можуть бути розгорнутими централізовано на одній віртуальній машині і розподілено на декількох віртуальних машинах. VNF також можна розділити на кілька підфункцій, так звані компоненти VNF (VNFC) під керуванням підсистеми EMS (Elemental Management Systems - EMSs);

2) NFV Infrastructure (NFVI). Інфраструктура NFV включає в себе все апаратне і програмне забезпечення, необхідне для розгортання, експлуатації та моніторингу функцій віртуалізованої мережі. З цією метою NFVI має рівень віртуалізації, необхідний для абстрагування апаратних ресурсів (обробка, зберігання і підключення до мережі). Це забезпечує незалежність програмного забезпечення VNF від фізичних ресурсів. Рівень віртуалізації зазвичай складається з серверних (Xen, KVM, VMware тощо) і мережевих (VXLAN, NVGRE, OpenFlow і т.п.) гіпервізорів. З боку безпосередньо самих функцій віртуалізованої мережі, апаратна частина архітектури NFV представляється як єдина обчислювальна платформа, що має здатність масштабуватись;
3) NFV Management and Orchestration (MANO) - підсистема управління функціями віртуалізованої мережі і оркестрації MANO складається з трьох компонентів:
   (1) диспетчер віртуалізованої інфраструктури (VIM), який управляє і контролює взаємодію VNF з фізичними ресурсами, що знаходяться під його контролем (наприклад, розподіл, звільнення і облік);
   (2) підсистема управління функціями віртуалізованої мережеві (VNFM), яка відповідає за управління життєвим циклом VNF (наприклад, ініціалізація, зупинка і завершення);
   (3) оркестратор NFV (NFVO), який є підсистемою управління та відповідальний за організацію програмних ресурсів задля управління інфраструктурою NFV. При цьому, однією з головних його функцій є – організація і управління послугами на NFVI.

Також варто відзначити елемент системи підтримки операцій і системи підтримки бізнесу (OSS/BSS). Цей елемент включає в себе сторонні системи управління і допомагає MANO у виконанні політик мережі, або автоматично, або вручну.

2.4 Програмно-конфігуровані мережі з розподіленими контролерами

Архітектура фізично розподілених контролерів (РК) стала абсолютно новим підходом до організації рівня управління в ПКМ. Її виникнення частково пов'язано з двома основними недоліками властивих програмованим мережам: використання єдиного централізованого контролера і проблема масштабованості. В ПКМ масштабованість відображає здатність контролера обробляти декілька запитів маршруту пересилання від комутаторів. Як відомо, контролер ПКМ має обмежений ресурс при обробці запитів. Щоб вирішити цю проблему, дослідники намагалися обмежити кількість запитів маршруту пересилки, які відправляються на контролер. Однак така стратегія базується на додаванні інтелектуальних функцій комутатору і тим самим порушує саму концепцію ПКМ.

Аналіз аспектів надійності обумовлений тим, що один контролер має проблему єдиної точки відмови. У разі виходу контролера з ладу комутатори втрачають здатність пересилати нові пакети, і, врешті-решт вся мережа виходить з ладу. Для вирішення цієї проблеми була запропонована часткова модифікація функцій комутатора OpenFlow, а також надання йому гібридних властивостей. Відмінною особливістю нового OpenFlow-hybrid комутатора є те, що при обриві з'єднання з контролером ПКМ, він може перейти в режим традиційного комутатора. Іншим рішенням може бути технологія фізичного РК, де на одному рівні управління використовуються кілька контролерів, які рівномірно розподіляють навантаження між собою.

У запропонованій раніше архітектурі програмно-конфігурованих мереж рівень управління використовував єдиний централізований контролер, який керує всіма процесами організації і контролю на декількох мережевих пристроях. В ході досліджень було виявлено, що дана фізично централізована архітектура може бути ефективною для малих мереж, однак великі мережі з високими вимогами не можуть працювати через один контролер, так як він представляє «вузьке місце» в мережі. Використання єдиного централізованого

контролера більше не відповідає нинішнім вимогам мереж, що спонукає проектувальників до застосування ПКМ з розподіленими контролерами в своїх проектах. Ідея даних мереж полягає у використанні кількох контролерів замість одного для управління платформою передачі даних.

Програмно-конфігуровані мережі с фізично розподіленими контролерами поділяються на дві основні категорії: логічно централізовані і логічно розподілені архітектури. Далі, логічно розподілену категорію можна розділити на дві підкатегорії: плоско розподілені і ієрархічно розподілені (рисунок 2.4).

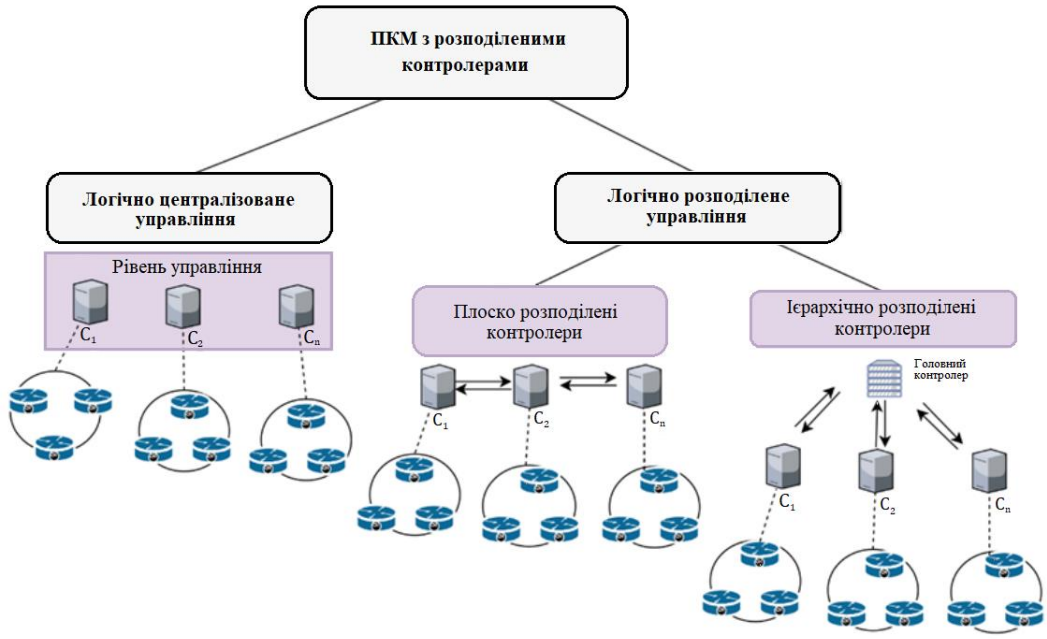

Рисунок 2.4 Архітектура ПКМ з фізично розподіленими контролерами

У логічно централізованій архітектурі обов'язки розподіляються порівну між кількома контролерами. Однак для рівня передачі даних вся структура виглядає як єдиний контролер, який керує всією мережею. Всі контролери завжди отримують повідомлення про будь-які зміни в мережі, використовують єдину базу даних і миттєво обмінюються інформацією завдяки мережевій синхронізації, завдяки мережевій синхронізації. Коротше кажучи, логічно централізована архітектура залишається близькою до початкової концепції ПКМ, яка використовує єдиний контролер.

В логічно розподіленої архітектурі контролери розподілені фізично і логічно. Крім того, кожен контролер має уявлення тільки про ті мережеві пристрої, за які він відповідає, і приймати рішення щодо конфігурації може лише для них.

Дана категорія має дві підкатегорії:

1) *Плоско розподілена архітектура*. У плоскій або горизонтальній архітектурі всі контролери розташовані на одному рівні управління, кожен контролер має однакові обов'язки, відповідає за свою частину мережі і взаємодіє з іншими на основі заздалегідь оголошеного механізму.
2) *Ієрархічно розподілена архітектура.* Ієрархічна архітектура використовує більше одного рівня контролерів. Це означає, що деякі контролери мають більше прав та обов'язків. Найефективнішим використанням таких мереж є мережа, в якій використовується головний контролер над рівнем розподілених контролерів. Логічно розподілена архітектура походить від першої концепції ПКМ, оскільки вона пропонує розподіл певних обов'язків між контролерами всередині мережі.

Крім даних категорій і підкатегорій програмно-конфігурованих мереж з розподіленими контролерами, вони відрізняються використанням статичного або динамічного алгоритму.

У статичному алгоритмі позиції контролерів і підключення до комутаторів визначаються під час конфігурації мережі і залишаються незмінними з часом.

У динамічному алгоритмі позиції контролерів і підключення до комутаторів можуть змінюватися з часом, що робить мережу гнучкою, дозволяючи адаптуватися до змін навантаження в мережі.

2.5 Алгоритм динамічного розподілу ПКМ-контролерів

Ієрархічна структура розподілених контролерів в програмно-конфігурованих мережах використовує кілька контролерів на рівні управління з механізмами динамічної кластеризації для балансування навантаження між ними.

Модельна мережа розглядається як загальна трирівнева система ПКМ з декількома контролерами в рівні управління і механізмами динамічної кластеризації для балансування навантаження між контролерами. Відповідно до даної структури, рівень управління складається з двох підрівнів: нижній, що містить розподілений контролер, і верхній, який представляє собою головний контролер (ГК). ГК бере на себе відповідальність за організацію і контроль РК, а також виступає в якості шлюзу для зв'язку з контролерами інших мереж (рисунок 2.5).

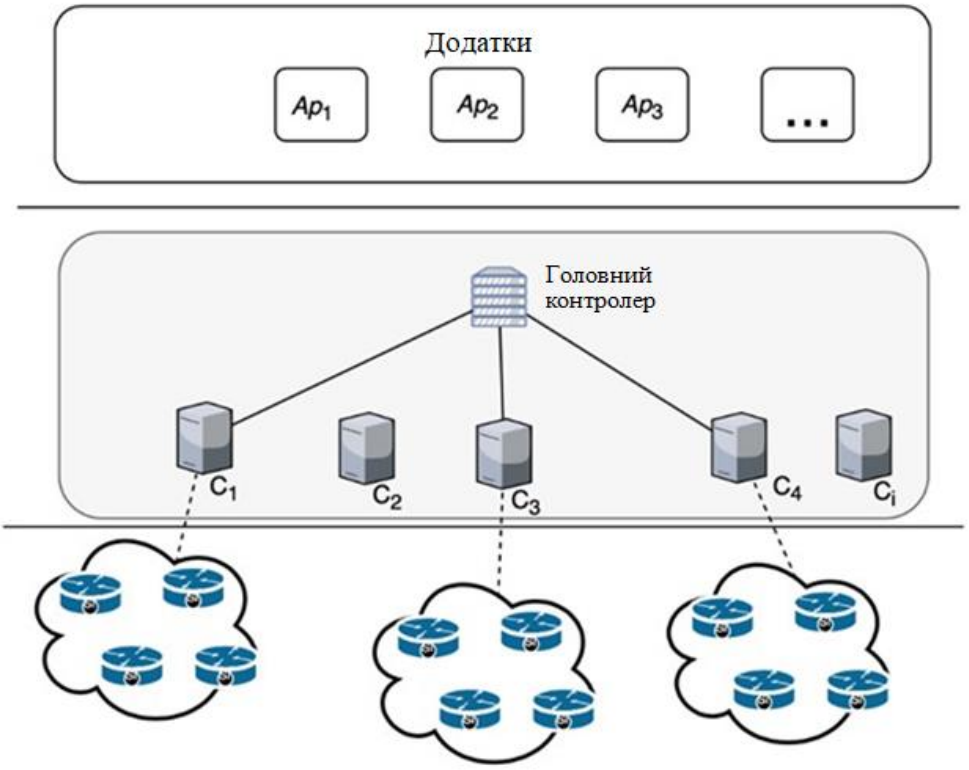

Рисунок 2.5 ПКМ як трирівнева система

РК організовані в кластери, кожен з яких містить лише один ПКМ-контролер і групу пристроїв передачі даних (комутатори OpenFlow). РК працюють разом під керуванням ГК для досягнення необхідної масштабованості і надійності мережі. ГК відповідає за налаштування розподілених кластерів та інформування кожного контролера про членів кластера - комутатори OpenFlow.

Формування кластера є динамічним процесом. Це означає, що кластери час від часу можуть змінюватися в залежності від стану контролера. Алгоритм DDC (Dynamic and Distributed Clustering, алгоритму динамічного розподілу) змінює кластери за допомогою робочого навантаження кожного контролера. ГК отримує періодичні звіти про робоче навантаженні і вирішує зберігати кластери або змінювати їх.

Вся робота запропонованої структури ПКМ ділиться на раунди, які з точки зору тривалості є гетерогенними. Кожен раунд складається з двох основних етапів: настройки і стійкого стану.

На етапі налаштування ГК формує кластери і оголошує контролеру комутатори OpenFlow, які є членами його кластера. ГК вибирає структуру кластера, грунтуючись на періодичних повідомленнях робочого навантаження кожного РК.

На етапі стійкого стану РК зв'язується з комутаторами і керує ними в своєму кластері. Набір комутаторів OpenFlow створює свої таблиці пересилання на основі інструкцій, отриманих від РК. Періодично кожен контролер перевіряє і обчислює свою робоче навантаження W_i, а потім відправляє звіт ГК, який порівнює робоче навантаження кожного контролера з максимальним робочим навантаженням W_{max}, призначеним для всіх контролерів. Передбачається, що всі РК є однорідними з точки зору апаратних і програмних можливостей і, таким чином, мають один і той же W_{max}. Потім ГК обчислює відсоткове навантаження L_i кожного контролера за формулою:

$$L_i = \frac{W_i}{W_{max}} * 100\%$$

ГК має два максимальних граничних значення для робочого навантаження контролера: жовтий і червоний. Якщо навантаження перевищує 80%, контролер долає перший жовтий поріг, і ГК поміщає його в таблицю завантажених контролерів, які можуть бути перевантажені, якщо навантаження і далі продовжуватиме зростати. Коли навантаження досягає 90% (червоний поріг), це означає, що контролер скоро буде перевантажений. ГК приймає рішення про перебудову кластера для подолання проблеми перевантаження, яка може привести до відмови або блокування контролера.

Для всіх РК в таблиці завантаження ГК організовує сусідні контролери, а також суміжні сусіднім контролерам на основі їх завантаження, включаючи їх в реорганізацію кластера. Якщо все сусідні контролери мають відсоткове навантаження нижче жовтого порога, ГК перебудовує кластер, що містить завантажений контролер, і сусідні, щоб подолати проблему можливого перевантаження мережі. Якщо один з сусідніх контролерів також завантажений, ГК перебудовує кластери цих сусідніх контролерів і суміжні сусіднім контролерам. При такій зміні робоче навантаження рівномірно розподіляється серед кластерів, і перевантаження контролера зникає.

2.6 Висновки до розділу 2

У другому розділі роботи детально розглядаються технології SDN та NFV. Тут описана передісторія, наведена архітектура та принципи побудови програмно-конфігурованих мереж. Після чого можна виділити 3 основних принципи ПКМ:

- поділ площини управління і передачі трафіку, в результаті чого комутатор обслуговує лише потік даних, а як наслідок стає більш простим та дешевим. Управління вище указаними комутаторами

проводить окремий SDN-контролер в якому знаходиться вся таблиця маршрутизації;

- централізоване управління, дозволяє використовувати єдиний, стандартний, відкритий інтерфейс між пристроями управління і передачі (OpenFlow), який, в свою чергу, знімає рамки вендорності;
- програмованість сервісів мережі, дана функція надає додаткам право запрошувати ресурси у мережі за вимогою.

Також в рамках цього розділу була розглянута технологія NFV та її архітектура. Вона віртуалізує фізичні елементи телекомунікаційної мережі шляхом виконання мережевих функцій програмними модулями, які працюють на стандартних серверах і віртуальних машинах в них, при чому дані модулі взаємодіють між собою як і їх апаратні рішення. Виходячи з вище сказаного мережеве обладнання може бути розташоване як і в ЦОД, так і в приміщеннях кінцевих користувачів.

В результаті роботи над розділом, виясняється, що існує проблема використання єдиного централізованого контролера, тому що він має обмежений ресурс при обробці запитів від комутаторів. При відмові контролера комутатори втрачають управління, що приводить до падіння всієї мережі. Це наштовхує проектувальників до впровадження ПКМ з розподіленими контролерами, ідея яких полягає у використанні кількох контролерів на рівні управління, замість одного, що дозволяє зменшити ризики відмови балансуючи навантаження між ними.

3 МОДЕЛЮВАННЯ ПРОГРАМНО-КОНФІГУРОВАНОЇ МЕРЕЖІ ДЛЯ ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ ЇЇ ФУНКЦІОНУВАННЯ

Для створення моделі мережі, а також динамічного моделювання її роботи, аналізу та оптимізації її характеристик, управління трафіком, звичайно, потрібно використовувати один з найпотужніших інструментів дослідження складних систем - імітаційне моделювання.

3.1 Загальне уявлення системи

В даний час для балансування навантаження в мережах зв'язку найчастіше використовуються апаратні рішення. Однак згідно з поточними прогнозами широке використання ПКМ для надання послуг зв'язку змінить процеси та методи балансування навантаження.

На рисунку 3.1 показаний кластер контролерів, підключених до планувальника завдань. Кожен з контролерів є елементом кластера, але працює незалежно від інших. Запити, які поступають від комутаторів надходять в буфер планувальника, який змінює адресу одержувача і пересилає запит до обраного контролера. Отримавши запит, контролер надсилає відповідь на адресу контрольного планувальника, який, в свою чергу, змінює адресу відправника на його віртуальну адресу і перенаправляє відповідь на комутатор одержувача.

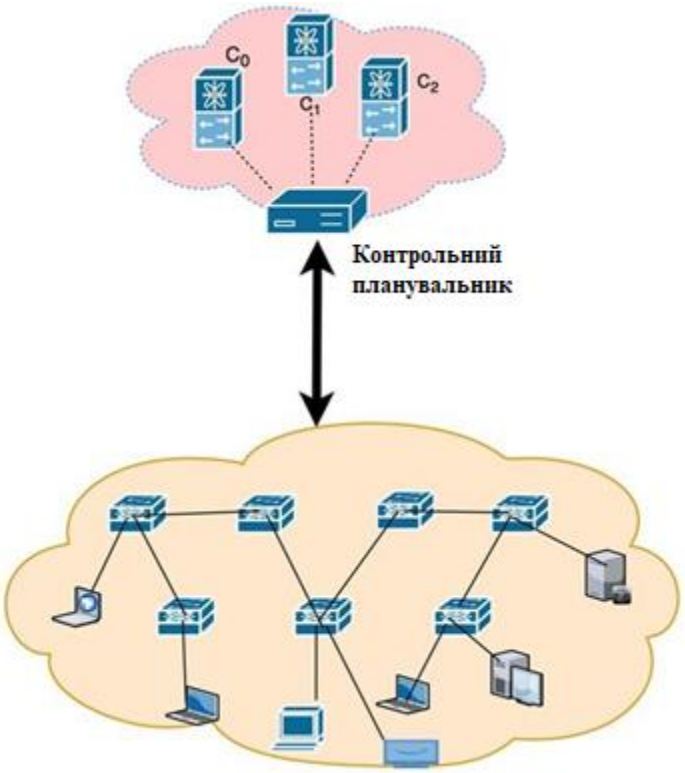

Рисунок 3.1 Модель мережі для кластеру контролерів

3.2 Імітаційна модель програмно-конфігурованої мережі

Традиційно в системах розподілу навантаження кілька взаємопов’язаних вузлів об’єднуються в один потужний вузол. Ці автономні вузли працюють незалежно один від одного, утворюючи команди для обслуговування та розподілу завдань. Щоб скористатися можливістю розподілу навантаження між усіма вузлами, що беруть участь в роботі, потрібні схеми взаємодії та оптимізовані алгоритми розподілу навантаження. Таким чином, вибір тієї чи іншої схеми обґрунтовується безліччю параметрів, таких як політика надсилання запиту, політика вибору обслуговуючого пристрою, а також клас вхідного запиту.

Розглянемо фрагмент ПКМ з декількома контролерами і представимо його як систему масового обслуговування з чергами, як показано на рисунку 3.2. Черга на вході контрольного планувальника формується відповідно до часу

прибуття. Далі в даному розділі будуть розглянуті дві моделі: аналітична і імітаційна.

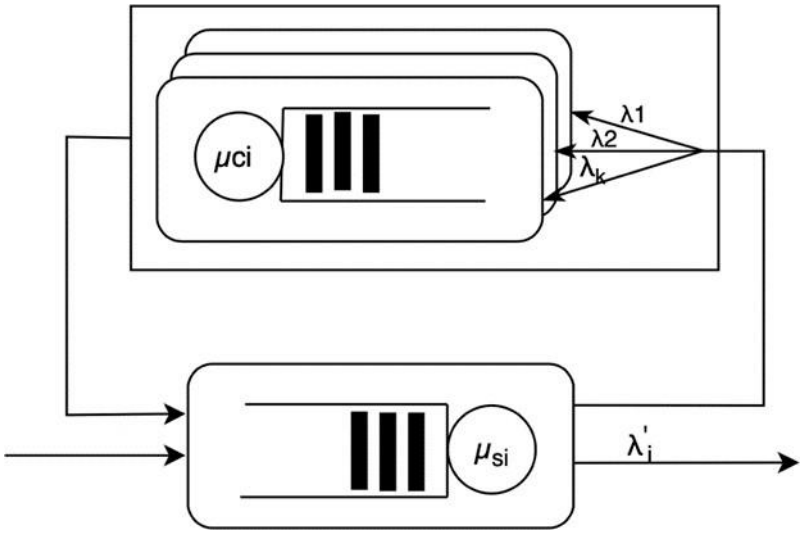

Рисунок 3.2 СМО для представленої системи контролерів

В ході дослідження були введені обмеження для спрощення реалізації фрагмента мережі, представленого на рисунку 3.1. Передбачається, що кожна заявка обслуговується незалежно одна від одної, а також може обслуговуватися будь-яким контролером кластера. В рамках дослідження розглядався лише процес розподілу навантаження, який безповоротно розподіляє заявки по контролерам, тобто заявка присвоюється до одному контролеру без повторного перерозподілу. Для простоти передбачається, що модель системи стійка і контролери ідентичні, тобто мають однакову швидкість обробки заявок. Імітаційна модель враховуватиме класи заявки і динамічно визначатиме стан контролера для обслуговування.

3.3 Аналітична модель програмно-конфігурованої мережі

При надходженні запиту на контрольний планувальник визначаються адреса і рівень завантаженості контролера кластера, після чого запит направляється на адресу мало завантаженого контролера, якщо такий існує, або на найменш завантажений з наявних. Динамічний процес розподілу і

присвоювання запитів контролерам є головним завданням для зменшення часу відгуку системи і ефективного використання його ресурсів.

Модель системи масового обслуговування для дослідження процесу обміну даними між комутаторами і кластером контролерів представлена на рисунку 3.2. Система складається з m обслуговуючих контролерів та k запитів що надходять на контрольний планувальник. У роботі передбачається, що модель роботи одного контролера може бути представлена як модель СМО (Система масового обслуговування) M/M/1, яку можна розширити до моделі M/M/m з метою аналізу можливостей функціонування ПКМ в разі використання кластера контролерів для розподілу навантаження. Передбачається, що запити надходять відповідно до закону розподілу Пуассона, а процес обслуговування запиту відбувається відповідно до експоненціального закону розподілу.

Для аналізу функціонування фрагмента мережі будемо використовувати значення середнього часу обслуговування запитів системи.

Нехай μ - інтенсивність обслуговування контролера і $\Delta\omega$ - необхідна якість обслуговування запиту. Для того, щоб гарантувати цю вимогу щодо якості обслуговування запиту, потрібно динамічно виділяти йому потрібні ресурси мережі. Загалом, μ є постійною величиною і, відповідно, щоб гарантувати необхідну якість обслуговування, запит повинен бути відправлений контролеру, здатному відповідати вимогам $\Delta\omega$. У зв'язку з цим, із збільшенням інтенсивності запитів може погіршитися якість обслуговування контролера, якщо кількість запитів перевищить його межу можливостей обслуговування. Таким чином, середній час обслуговування запиту змінюється в залежності від зміни інтенсивності і кількості запитів. Уявімо μ як $\omega(t)$ - час обслуговування контролером одного запиту в заданому інтервалі часу t.

В даному випадку, для представлення потрібної якості обслуговування, потрібно щоб ($\omega(t) \leq \Delta\omega$) для даної точки часу t.

Нехай p_k – ймовірність того, щоб в системі були присутні k запитів.

Тоді

$$p_k = p_0 \prod_{i=0}^{k-1} \frac{\lambda_i}{\mu_i}, p_0 = 1 - \frac{\lambda}{\mu}.$$

Система СМО M/M/1 може бути охарактеризована коефіцієнтами інтенсивності надходження і обслуговування запитів, такими як

$$\lambda_i = \lambda, k = 1,2,3 \dots, \mu_k = \mu, k = 1,2,3 \dots .$$

Застосовуючи ці коефіцієнти в p_k, отримаємо

$$p_k = p_0 \left(\frac{\lambda}{\mu}\right)^k, k \geq 0.$$

Для стійкості системи $0 \leq \rho < 1$, потрібно щоб $p_0 > 0$ і звідси можна сказати, що p_0 є постійною величиною. Відповідно, отримаємо

$$p_k = (1 - \rho)\rho^k, \rho = \frac{\lambda}{\mu}.$$

Застосовуючи закон Літтла $N = \lambda\omega$, можна розрахувати кількість запитів в чергу на обслуговування системи

$$N = \sum_{k=0}^{\infty} k p_k = \frac{\rho}{1 - \rho}.$$

Таким чином, отримаємо

$$\omega(t) = \frac{\frac{1}{\mu}}{1 - \rho} = \frac{1}{\mu - \lambda}. \#(1)$$

Для рівняння (1) передбачається, що $\omega(t)$ відповідає очікуванням щодо якості обслуговування запиту. Однак при збільшенні інтенсивності запитів, що надходять потрібні додаткові ресурси. Рівняння (1) не може обумовлювати процес обслуговування, тому застосовується модель M/M/m для аналізу середнього часу обслуговування контролера.

Передбачаються також система з нескінченним місцем у черзі і постійною інтенсивністю надходження запитів λ.

Система забезпечує максимальне m число контролерів, тому

$$\mu_k = \min[k\mu, m\mu] = \begin{cases} k\mu, 0 \le k \le m \\ m\mu, \quad m \le k \end{cases}$$

Для стійкості системи

$$\lambda = \frac{\lambda}{m\mu} \le 1.$$

Відповідно

$$p_k = \begin{cases} p_0 \dfrac{(m\rho)^k}{k!}, k \le m \\ p_0 \dfrac{\rho^k m^m}{m!}, k \ge m \end{cases}$$

А також для p_0 отримаємо

$$p_0 = \left[1 + \sum_{k=1}^{m-1} \frac{(m\rho)^k}{k!} + \sum_{k=m}^{\infty} \frac{(m\rho)^k}{k!} \frac{1}{m^{k-m}}\right]^{-1}$$

Імовірність того, щоб запит який надходить потрапив в чергу

$$p_q = \sum_{k=m}^{\infty} p_k = p_0 \frac{(m\rho)^m}{m!\,(1-\rho)} = 1 - \sum_{k=0}^{m} p_0 \frac{(m\rho)^k}{k!}$$

Так як нам потрібно знайти середній час обслуговування запиту $\omega(t)$, можна відзначити, що число m контролерів є ще і параметром $\omega(t)$.

Відповідно,

$$\omega(t,m) = \mathrm{E}[\omega(t,m)] = \frac{1}{\lambda}(m\rho + \rho \frac{(m\rho)^m}{m!} \frac{p_0}{(1-\rho)^2})$$

Застосовуючи $\lambda = \frac{\lambda}{m\mu}$ в (1) отримаємо

$$\omega(t,m) = \frac{1}{\mu} + \frac{1}{\lambda} \frac{\left(\frac{\lambda}{\mu}\right)^m}{m!} \frac{p_0}{\left(1 - \frac{\lambda}{m\mu}\right)^2}. \#(2)$$

Як було розглянуто вище, для того, щоб гарантувати необхідну якість обслуговування, нерівність (3) має бути дотримана в такий спосіб:

$$\frac{1}{\mu} + \frac{1}{\lambda} \frac{\left(\frac{\lambda}{\mu}\right)^m}{m!} \frac{p_0}{\left(1 - \frac{\lambda}{m\mu}\right)^2} \leq \frac{1}{\mu - \lambda}. \#(3)$$

Нехай $r \geq 1$ визначає додатковий трафік в системі, тоді, множачи λ на r і спрощуючи рівняння (2) (при цьому повинні бути задоволені $r\lambda/\mu \leq m$ і $r \geq 1$), отримаємо

$$f(r,m) = \frac{\lambda}{\mu} - (\mu - \lambda)\frac{(r\lambda)^{m-1}}{m!\,\mu^m}\frac{p_0}{\left(1 - \frac{r\lambda}{m\mu}\right)^2}$$

Модель працює за наступним алгоритмом планування розподілу:

1) визначити кількість контролерів m в кластері С;
2) для всіх контролерів визначити інтенсивність обслуговування: μ_k, $\forall k$;
3) визначити інтенсивність надходження запитів на контрольний планувальник в інтервал часу *t:* $\lambda_k(t)$;
4) для всіх контролерів визначити $\omega(t, m)$,
5) визначити контролер c_{target} для обслуговування запиту який надходить:
   - задати проміжне максимальне значення часу обслуговування запиту;
   - перевірити з кластера С час відгуку ω_j контролера c_j:
   - якщо $\omega_j < \omega_{jmax}$, то $\omega(t, m) = \omega_j$ і $c_{target} = c_j$;
   - для $j = 1, ..., m$ пройти весь цикл і знайти менш завантажений контролер c_{target} з найменшим часом обслуговування;
   - якщо c_{target} не знайдеться, зберігати запити в чергу;
6) під час надходження нового запиту, повторювати процес з кроку 3.

3.4 Результати моделювання

Для оцінки запропонованого алгоритму була розроблена імітаційна модель фрагмента ПКМ в середовищі AnyLogic. Модель описує процес обслуговування вхідних повідомлень на різні блоки обслуговування.

Розглянувши аналітичний метод оцінки якості обслуговування в мережі ПКМ, ми проаналізували ситуацію в разі зміни трафіку. Для зручності було розглянуто випадок використання одного контролера із збільшеною продуктивністю в 3 рази і проведено порівняння з випадком використання кластера з 3 контролерів з однаковими значеннями продуктивності. Результати моделювання представлені в малюнках нижче.

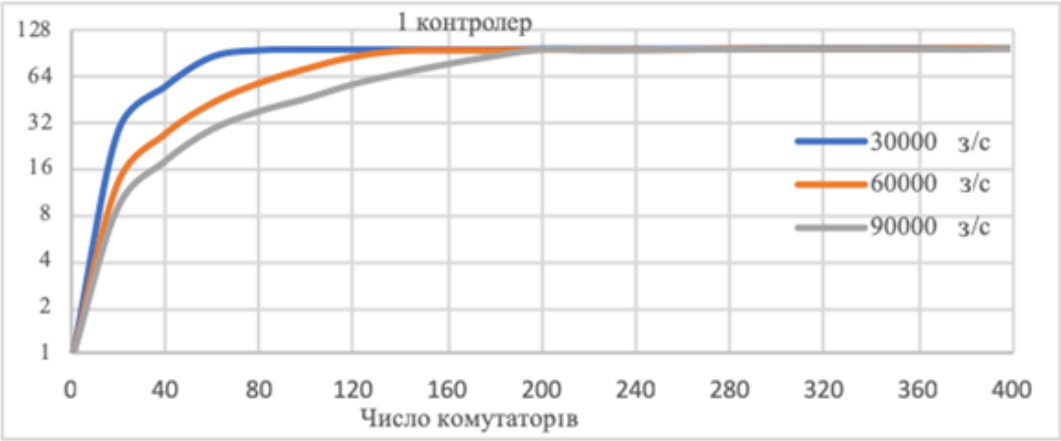

Рисунок 3.3 Залежності навантаження контролера від кількості підключених комутаторів в разі одного контролера

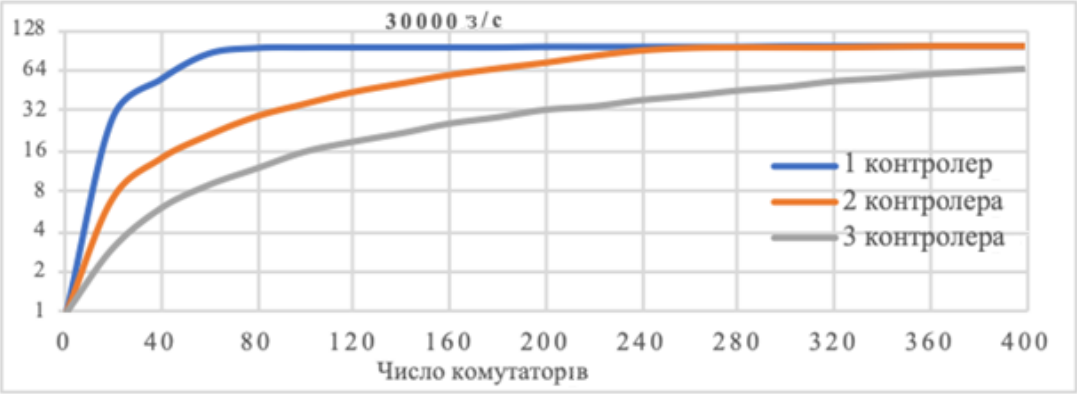

Рисунок 3.4 Залежності навантаження контролера від кількості підключених комутаторів в разі кластера з 3 контролерів

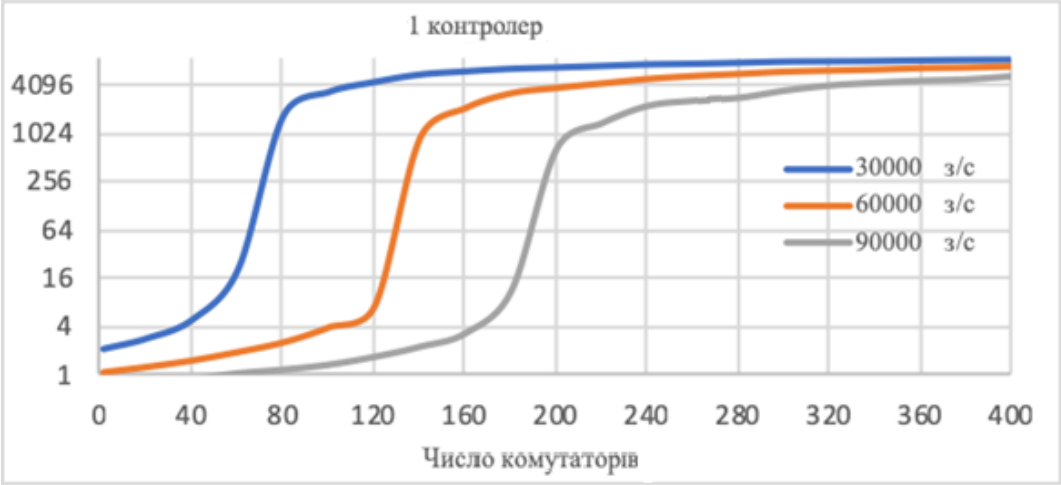

Рисунок 3.5 Залежності середнього часу обслуговування контролера від кількості підключених комутаторів в разі одного контролера

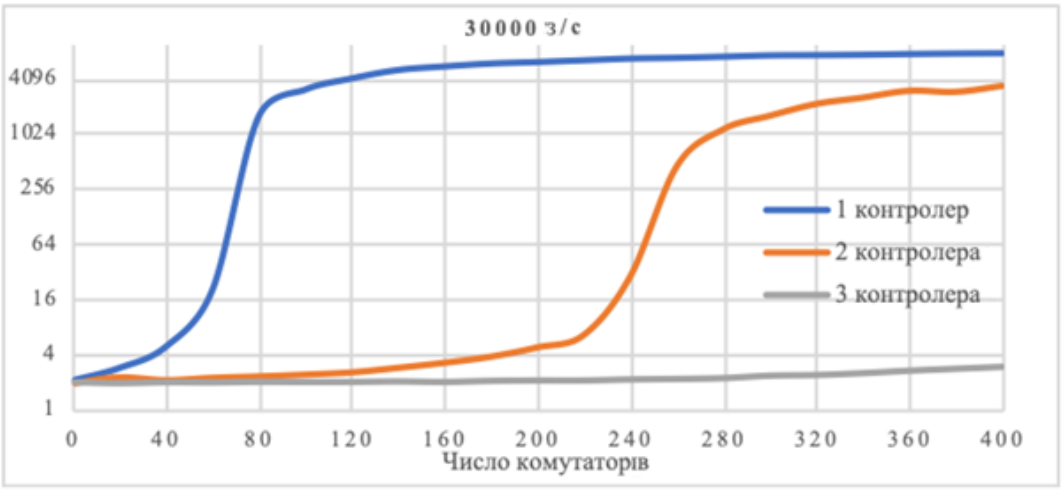

Рисунок 3.6 Залежності середнього часу обслуговування контролера від кількості підключених комутаторів в разі кластера з 3 контролерів

На рисунках 3.3 і 3.4 показані зміни навантаження контролера, в залежності від кількості підключених комутаторів. Графік залежності показує, що зі збільшенням числа комутаторів відповідно збільшується навантаження на контролері аж до критичної позначки. Пропорційно збільшується середній час обслуговування запиту, що показано на рисунку 3.4.

На рисунках 3.3 і 3.5 показані графіки результатів моделювання системи з одним контролером, при цьому були встановлені значення продуктивності ядра (μ) в 30000 з/с, 60000 з/с і 90000 з/с відповідно. Результати показують, що контролер може обслуговувати, в кращому випадку, до 180 комутаторів до того, як погіршиться показник обслуговування, обумовлений середнім часом обслуговування. На рисунках 3.4 і 3.6 наведені графіки результатів моделювання системи, де був використаний алгоритм планування розподілу навантаження; замість одного контролера з продуктивністю μ = 90000 взяли кластер з 3 контролерів, кожен з продуктивністю μ = 30000 з/с. Результати показують, що система може обслуговувати до 500 комутаторів без істотного зменшення показників обслуговування.

Проведений аналіз показує, що використання кластера контролерів в 3 рази покращує можливості мережі в забезпеченні необхідної якості обслуговування. Це пояснюється тим, що контролери можуть

використовуватися раціонально та залежно від потреб мережі вони можуть перебувати в різних режимах (очікування, сплячий), що покращує процес балансування навантаження.

3.5 Розробка моделі класифікації і пріоритезації трафіку в програмно-конфігурованих мережах

3.5.1 Характеристики мережевого трафіку

Під'єднані до мережі додатки генерують трафік (вхідний та вихідний), ґрунтуючись на конкретні особливості конфігурації програми. Передані пакети включають трафік конфігурації мережі, протокол мережевого часу (NTP, від англ. Network Time Protocol), систему доменних імен (DNS, від англ. Domain Name System), зв'язок між пристроями і сервером, та трафік, що генерується при взаємодії з користувачем.

Хоча різні додатки в мережі можуть використовувати різні протоколи та передавати дані для різних цілей, більшість трафіку використовує протоколи TCP/IP (Transfer Control Protocol/ Internet Protocol). Для того, щоб однозначно визначити, чи належить пакет до певного потоку, в полях MatchField таблиці OpenFlow вказуються відповідні значення. Таким чином, виходячи з групи параметрів, наприклад, IP-адреса джерела/призначення можна вибрати відповідний потік і моніторити показники лічильника потоку (Packet count, Byte count).

Кожен потік трафіку містить основну інформацію про пакет, від MAC (Media Access Control)-рівня до рівня додатків. Мережевий трафік може розглядатися як дані тимчасових рядів, що містить корисну інформацію про користувачів, пристрої і стан мережі. У цьому випадку для збору трафіку використовувались аналізатори пакетів трафіку Wireshark. Через обмеження засобів мережевої безпеки, таких як протокол рівня захищених сокетів (SSL, від англ. Secure Sockets Layer) і протокол захисту транспортного рівня (TLS, від

англ. Transport Layer Security), для класифікації можна використовувати лише заголовки пакетів.

Відповідно, для класифікації мережевого трафіку необхідний оптимальний набір ознак. Ознака трафіку - це атрибут, який має різне значення для різних типів класів трафіку. Наприклад, середній розмір пакета, зазвичай, різний для потоків мультимедійного контенту і потоків завантаження, оскільки в останніх майже всі пакети є повнорозмірними, що не стосується мультимедійних потоків.

3.5.2 Модифікований алгоритм кластеризації k-means

Метою алгоритму є вирішення проблеми кластеризації потоків трафіку в мережі зв'язку [34]. Передбачається, що один потік трафіку реалізує (або бере участь в реалізації) однієї з можливих послуг зв'язку. У цьому випадку можливий кінцевий набір k видів трафіку, наприклад, передача відео, передача музики, мови, інтерактивного відео, завантаження файлів і ін. Кожен з видів трафіку має певні характеристики, які відображаються в його параметрах, можливо, в деяких ознаках, отриманих шляхом його моніторингу. Нехай d - кількість таких характеристик. При виконанні практичних експериментів $d = 13$. Перелік параметрів наведено в Таблиці 3.1.

Візьмемо за основу алгоритм кластеризації k-means [35], який дозволяє виділяти задану кількість кластерів. Модифікуємо цей алгоритм з метою його застосування в даному випадку. Особливість кластеризації (класифікації) потоків полягає в наступному: загальна кількість характеристик трафіку, доступних для моніторингу досить велике; трафік характеризується різними параметрами з різними одиницями виміру і діапазонами можливих числових значень; кількість спостережень (результатів моніторингу, потоків) змінюється з часом.

Таблиця 3.1 – Набір ознак для моделі класифікації трафіку

Назва	*Опис*
Source IP (src_IP)	IP адреса джерела
Destination IP (dst_IP)	IP адреса призначення
Source Port (src_port)	Порт джерела
Destination Port (dst_port)	Порт призначення
Average window size	Середній розмір розглянутого набору потоків, *байт*
Number of packets	Кількість пакетів
Packet size	Розмір пакетів, *байт*
Average packet size	Середній розмір пакета, *байт*
Standard deviation of packet sizes	Стандартне відхилення розміру пакетів, *байт*
Average inter-arrival time	Середній час надходження пакетів, *с*
Standard deviation of inter-arrival times	Стандартне відхилення часу надходження пакетів, *мс*
Flow duration	Тривалість потоку, *с*
Flow size	Розмір потоку, *байт*

Розглянемо d-мірний простір, в якому координати точки (елемента) визначаються d числами. Будемо вважати, що простір розглянутих характеристик трафіку є метричним. Відстань між двома точками x_i і x_j (за точку може бути прийнятий потік по результату його моніторингу) визначається як

$$S(i,j) = \sqrt{\sum_{r=1}^{d} \left(x_i^{(r)} - x_j^{(r)} \right)^r} . \#(4)$$

Будемо вважати, що значення характеристик потоків (параметрів) можуть змінюватися від деякого мінімального до деякого максимального значення

$$c_{min}^{(r)} \le x^{(r)} \le c_{max}^{(r)}, r = 1 \dots d. \#(5)$$

Оскільки характеристики потоку можуть мати різні одиниці вимірювання і різні діапазони можливих значень, будемо нормувати їх значення

$$\tilde{x}^{(r)} = \frac{1}{c_{max}^{(r)}}\left(\tilde{x}^{(r)} - c_{min}^{(r)}\right), r = 1 \dots d. \#(6)$$

Тоді

$$0 \le \tilde{x}^{(r)} \le 1, r = 1 \dots d. \#(7)$$

Робота алгоритму складається з двох основних процесів: «навчання» (або адаптація) і класифікація потоків. Навчання полягає у виділенні заданої кількості k кластерів і обчисленні їх центрів мас, тобто координат центрів кластерів, як

$$x_{cm}^{(r)} = \frac{1}{m_r}\sum_{i=1}^{m_r} \tilde{x}_i \,, r = 1 \dots d. \#(8)$$

де $\tilde{x}$ нормоване, згідно (7), значення r-й характеристики потоку.

Виділення кластерів проводиться відповідно до алгоритму k-means, тобто являє собою ітераційну процедуру, в ході якої проводиться перерозподіл елементів по кластерам і перерахунок центрів мас, поки центри кластерів не стабілізуються.

Знайдені центри мас можуть бути використані в задачі класифікації потоків трафіку. Дане завдання може вирішуватися оцінкою ступеня близькості даного потоку до центрів мас згідно

$$S(i,w) = \sqrt{\sum_{r=1}^{d} \left(x_i^{(r)} - x_{cm.w}^{(r)}\right)^2}. \ \#(9)$$

Належність даного потоку деякого типу потоків (кластеру) може бути визначено як

$$\hat{r} = argmin_w S(i,w). \ \#(10)$$

На відміну від «класичного» алгоритму, в даному випадку кількість об'єктів (потоків), що підлягають кластеризації змінюється в часі, тобто збільшується в процесі виконання моніторингу. На початку спостережень кількість об'єктів мало, і результат кластеризації може бути недостовірним.

Для оцінки отриманого результату обчислюються середньоквадратичні відхилення елементів кластерів від їх центрів мас

$$\sigma_w = \sqrt{\frac{1}{m_w - 1}\sum_{i=1}^{m_w} S(i,w)^2}. \ \#(11)$$

де $S\ (i,\ w)$ - відстань між *i*-м елементом *w*-го кластера і центром мас *w*-го кластера, згідно (9), m_w - кількість елементів в *w*-м кластері.

Також обчислюється загальне середньоквадратичне відхилення для всіх елементів

$$\sigma = \sqrt{\frac{1}{n - 1}\sum_{i=1}^{n} S(i,x_0)^2}, \ \#(12)$$

де *n* - загальна кількість елементів;

$S(i,\ x_0)$ - відстань між *i*-м елементом і загальним центром мас x_0.

$$x_0^{(r)} = \frac{1}{n}\sum_{i=1}^{n} x_i^{(r)}, r = 1 \dots d, \#(13)$$

Порівняння σ_w і σ дозволяє судити про якість виконання завдання кластеризації потоків. Чим менше величина

$$\delta_w = \frac{\sigma_w}{\sigma}, w = 1 \dots k, \#(14)$$

тим менше розкид елементів всередині кластера w, в порівнянні з розкидом між усіма елементами без поділу на кластери.

Таким чином, за допомогою (11) і (14) можна характеризувати рішення про віднесення потоку до деякого кластеру (типу).

Доцільно ввести деяке порогове значення δ_0, яке свідчить про можливість прийняття рішення. Інакше кажучи, рішення про віднесення потоку до деякого типу w може бути прийнято тільки тоді, коли $\delta_w \leq \delta_{w0}$, $w = 1 \dots k$. Величина порогового значення може вибиратися емпірично, на основі зібраних даних моніторингу.

Порівняння σ_w с $S\ (i,\ w)$ дозволяє оцінити ступінь близькості i-го потоку до потокам обраної групи. Чим менше величина

$$\eta_{w,i} = \frac{S(i,w)}{\sigma_w}. \#(15)$$

тим більше впевненість, що i-й потік відноситься до типу w. Згідно з правилом 3σ можна сказати, що якщо ця величина менше 1/3, то ймовірність того, що потік відноситься до типу w - не менше 0,99. Однак на практиці такі оцінки не завжди застосовуються, тому для цієї величини також доцільно вибрати деяке емпіричне значення η_0 і приймати рішення, якщо $\eta_{w,i} \leq \eta_0$.

Таким чином, модифікація алгоритму k-means полягає у визначенні розмірності простору, правил оцінки чисельних характеристик і способу оцінки якості прийнятого рішення.

Ефективність даного методу в порівнянні з «класичним» алгоритмом складається в можливості обліку різних характеристик трафіку і оцінки якості рішення, що «класичний» алгоритм не дозволяє зробити. Це виражається в зниженні помилки прийняття помилкових рішень класифікації потоків.

3.5.3 Модель класифікації та пріоритезації трафіку ПКМ

Модельна мережа складається з ПКМ-додатку, який класифікує мережевий трафік і приймає рішення про пріоритезацію трафіку, клієнтських агентів (хост-пристроїв) з додатками, які генерують мережевий трафік і маршрутизатори, які застосовують правила пріоритетів трафіку до активних потоків (рисунок 3.7).

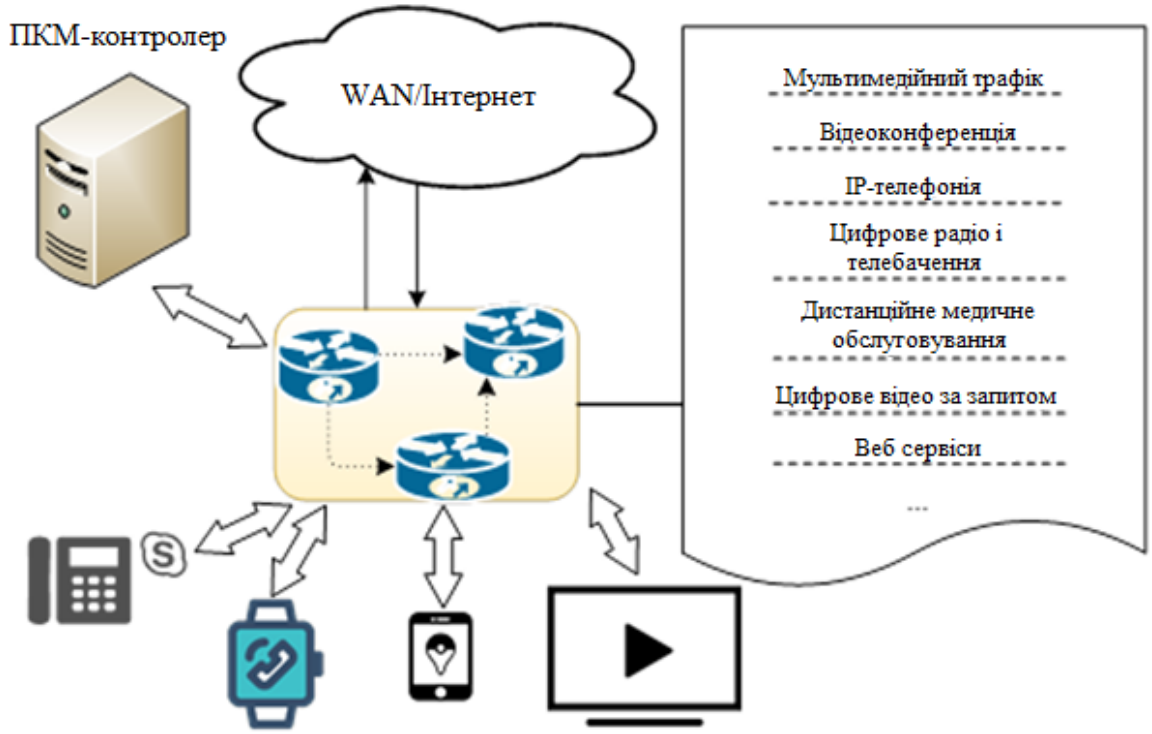

Рисунок 3.7 Модель ПКМ для завдання класифікації трафіку

Запропонований метод класифікації та пріоритезації трафіку в ПКМ працює наступним чином. Для класифікації спочатку компілюється набір даних, з якого витягуються особливості потоків певного типу трафіку. Потім вибирається мінімальний набір ознак, які з високою точністю характерні для потоку; після чого застосовується алгоритм класифікації для навчання класифікатора, який в подальшому використовується в сценарії в реальному часі.

Оскільки одні й ті ж «навчальні» дані можуть бути неефективними через тривалий проміжок часу, оскільки характеристики потоку змінюються протягом певного періоду часу, необхідно регулярно оновлювати базу даних, тобто перекваліфікувати класифікатор. Таким чином можна з точно визначити зміни в характеристиках трафіку. Механізм динамічної класифікації схематично наведено на рисунку 3.8.

Рисунок 3.8 Механізм класифікації та пріоритезації трафіку

У процесі класифікації в режимі реального часу кожен потік класифікується як медіапотік або потік завантаження, триває додавання значень ознак і вектора класу в тимчасовий файл, і лише після того, як буде класифіковано певну кількість потоків, отримані значення додаються до набору ознак в існуючих даних для перекваліфікації моделі. На рисунку 3.9 представлена «теплова карта» кореляцій ознак потоку, де поле class представляє клас, до якого класифікується потік, у цьому контексті клас мультимедіа або клас завантаження.

Взаємозв'язок між очікуваними значеннями ознак трафіку і реальними для класифікації підтверджує ефективність роботи запропонованого методу, де для розглянутого мультимедійного трафіку і трафіку завантаження точність

класифікації потоку становить близько 98%. Ця карта кореляцій ознак потоку є критично важливою для для завдання визначення пріоритетності класифікованого типу трафіку. У разі неправильної класифікації потоку, відповідно, помилково до них будуть застосовані правила пріоритезації.

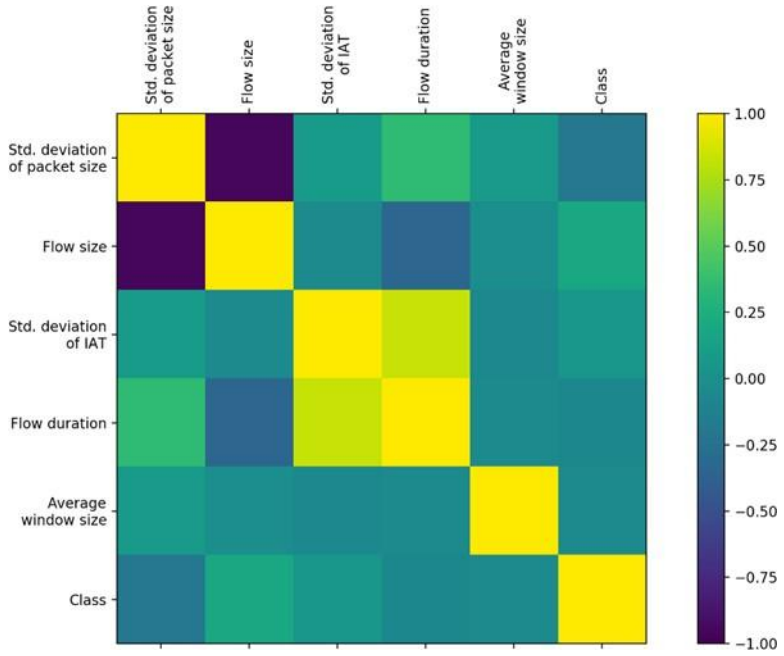

Рисунок 3.9 «Теплова» карта кореляцій ознак потоку

3.6 Висновки до розділу 3

В даному розділі дипломної роботи досліджувалась ефективність функціонування кластеру контролерів ПКМ та розроблювався алгоритм кластеризації для потоків трафіку мережі зв’язку.

Метою алгоритму являється групування потоків трафіку за певними ознаками для визначення приналежності потоку до певного кластеру. Що надалі дозволить налаштувати кластер під певну послугу та, наприклад, удосконалювати процес зняття плати.

Моделювання кластеру контролерів ПКМ проводилося для дослідження ефективності функціонування одного контролера в порівнянні кластером 3 контролерів. Розроблюється аналітична модель мережі, а потім на її основі

створюється імітаційна модель фрагмента ПКМ в середовищі AnyLogic, яка описує процес обслуговування вхідних повідомлень на різні блоки обслуговування.

Результати показують, що один контролер може обслуговувати до 180 комутаторів до того, як погіршиться середній час обслуговування заявки. В свою чергу кластер з 3 контролерів з розподілом навантаження може обслуговувати до 500 комутаторів без значного погіршення показника обслуговування, який описується середнім часом обслуговування заявки. Отже проведений експеримент показує, що використання кластеру контролерів в 3 рази покращує можливості мережі в забезпеченні необхідної продуктивності.

4 РОЗРОБКА МОДЕЛІ ІДЕНТИФІКАЦІЇ ТА ПРИОРИТЕЗАЦІЇ ТРАФІКУ ІНТЕРНЕТУ РЕЧЕЙ НА ОСНОВІ СЕГМЕНТАЦІЇ РЕСУРСІВ

4.1 Концепція мережевої сегментації

Майбутні мобільні мережі зв'язку повинні спрощувати процес налаштування мережі відповідно до вимог якості конкретних послуг, що надаються, зокрема, встановлення необхідної швидкості передачі даних, затримки, надійності, безпеки та інших послуг для різних категорій користувачів. Для цього в мережах зв'язку 5G введена технологія сегментації. Концепція мережевої сегментації була раніше запропонована в контекстах архітектур розподілених сервісів і додатків; але його використання в мережах мобільного зв'язку є новим. З точки зору оператора мобільного зв'язку, мережевий сегментації - це процес створення за потребою користувача, набір логічно незалежних мереж (мережні сегменти), розташованих на загальній фізичній інфраструктурі; кожна з яких налаштована на представлення певних мережевих послуг (рисунок 4.1).

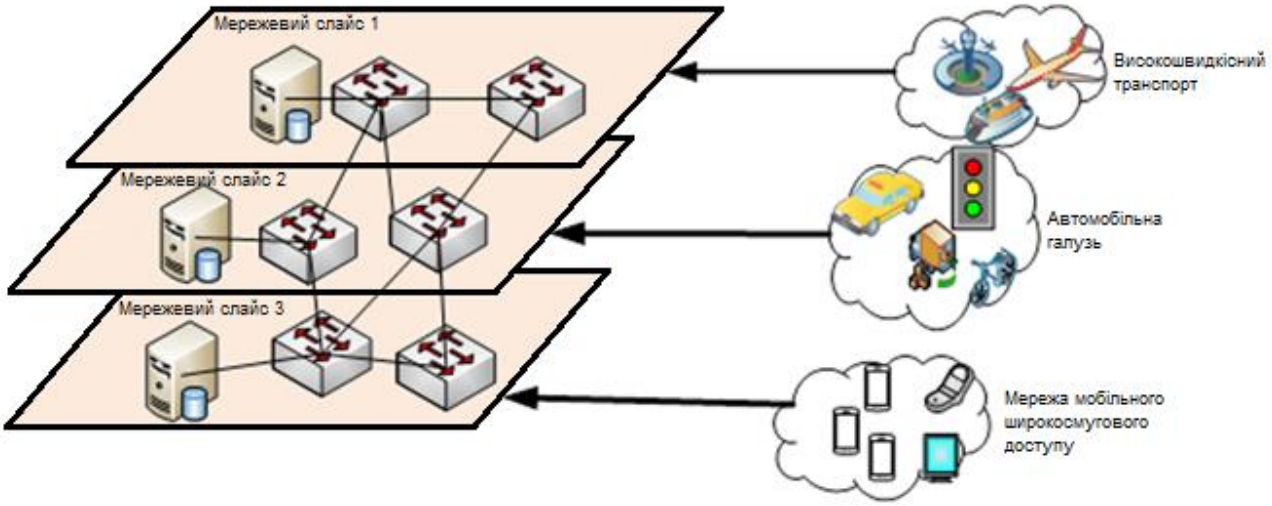

Рисунок 4.1 Архітектура мережевої сегментації в мережах зв'язку 5G

В контексті програмованих мереж, сегмент може бути представлений як віртуальна мережа, що працює незалежно від своєї фізичної інфраструктури.

Наприклад, віртуальний мережевий міст між віртуальним вузлом мережі і його фізичним вузлом. Віртуальний вузол мережі може здійснювати певні мережеві послуги як звичайний фізичний вузол (маршрутизатор, брандмауер, або сервер мережевих послуг). Встановлення віртуальних мережевих мостів може бути здійснено ПКМ-комутатором. Технологія SDN дозволяє адміністратору управляти фізичною мережею для того, щоб виділити необхідні ресурси створеному мережевого сегменту. А віртуальний мережевий вузол може бути встановлений при використанні технології віртуалізації мережевих функцій. Технології SDN/NFV дозволяють мережевій сегментації задовольняти вимогам по високому ступені гнучкості мережі.

4.2 Розробка моделі ідентифікації та пріоритезації трафіку Інтернету речей на основі сегментації ресурсів в ПКМ

4.2.1 Модель контролю параметрів якості обслуговування для пріоритезації додатків Інтернету Речей в ПКМ

Організація динамічної пріоритезації і управління трафіком додатків Інтернету Речей в умовах неоднорідності мереж дозволить впровадити такі нові послуги, як тактильний інтернет, доповнена реальність, медичні програми і інші.

Для забезпечення цих умов необхідна ініціалізація та ідентифікація пристроїв IoT (Internet of Things) із передачею вимог QoS від цих пристроїв до системи управління мережею. Слід також мати на увазі, що при побудові повноцінної інфраструктури мережі 5G буде існувати більше одного сегменту ПКМ, також будуть активно застосовуватися віртуальні сегменти (з одною або більше віртуальними мережевими функціями), для їх поєднання використовуються так звані оркестратори, які вже дозволяють працювати з підконтрольною інфраструктурою, як з єдиним ресурсом на більш високому рівні абстракції. Тому існує необхідність забезпечити процес передачі вимог QoS до додатків оркестратора мережі. Для цього була розроблена модель

протоколу, метою якого є організація процесу реєстрації Інтернет Речей та подальшого обслуговування за умови їх мобільності. Крім того, на модель покладено функції регулювання параметрів QoS шляхом організації роботи з оркестратором мережі, який в свою чергу працює вже з модулями ПКМ-контролера(-ів) (версія протоколу OpenFlow не менше 1.3), що відповідають за ці завдання в мережі і гіпервізорами (можливо також і з віртуальними або логічними сутностями) сегментів віртуальних мереж. Даний фреймворк реалізує певний алгоритм взаємодії систем (і/або їх елементів) в мережах зв'язку п'ятого покоління, тоді як взаємодія між деякими з них може відбуватися за допомогою різних протоколів. З появою та поступовим впровадженням різних сервісів IoT на працюючу мережу, запропонована модель дозволить диференціювати трафік різних додатків IoT та зможе забезпечити необхідні показники якості обслуговування для певних програм.

4.2.2 Архітектура високого рівня моделі та загальні описи взаємодії елементів

На рисунку 4.2 мережа зв'язку логічно розділена на кілька сегментів: рівень доступу, рівень агрегації, рівень ядра мережі, при цьому такий розподіл характерний для кожного з операторів.

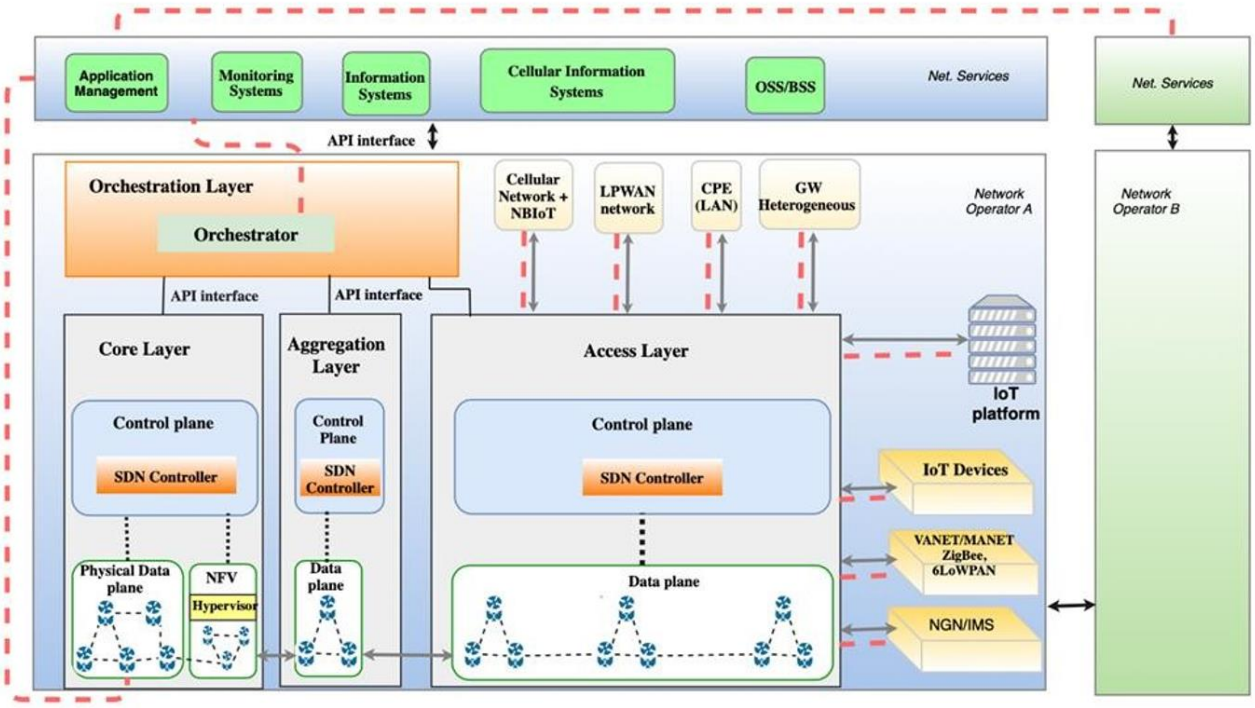

Рисунок 4.2 Інфраструктура мережі зв'язку з системою контролю QoS

У той же час кожним із сегментів мережі (доступ, агрегація, ядро) керує ПКМ-контролер, також можлива наявність і віртуальних сегментів мережі, кожен з яких може реалізовувати одну або більше мережевих функцій. Для реалізації взаємодії між сегментами потрібен мережевий оркестратор, який, у свою чергу, по південному інтерфейсу управляє контролерами і взаємодіє з гіпервізорами віртуальних сегментів мережі, а також безпосередньо з віртуальними машинами, які реалізують ту чи іншу мережеву функцію. Однак, враховуючи той факт, що більшість комерційних компаній з великими розподіленими корпоративними мережами вже переходять на технології програмно-конфігурованих мереж і віртуалізації мережевих функції, з упевненістю можна стверджувати, що дана модель дозволить реалізувати різноманітні додатки Інтернету Речей, які наразі важко реалізовані через ті вимоги, які вони пред'являють до мереж зв'язку. Наприклад, такі додатки як: доповнена віртуальна реальність, тактильний інтернет і медичні мережі дуже вимогливі до такого параметру, як час затримки, а деякі з них також і до швидкості, і її сталості в процесі надання послуги. Тому застосування цієї

моделі можливо спочатку в мережах корпоративного масштабу з подальшим перетворенням на комерційні мережі оператора.

На рисунку 4.2 червоною пунктирною лінією відображений взаємозв'язок між основними елементами. Наступні елементи є основними:

1) підтримка запропонованої моделі у вигляді програми оркестратора мережі (віртуальний або фізичний сервер);
2) підтримка даної моделі у вигляді бібліотеки мови розробки для пристроїв IoT, або іншого іншого блоку програмного забезпечення, що реалізує необхідний набір функцій і логіки реалізації взаємодії з іншими елементами моделі;
3) підтримка даної моделі у вигляді бібліотеки для сторонньої платформи IoT, або іншого блоку програмного забезпечення, що реалізує необхідний набір функцій і логіки реалізації взаємодії з іншими елементами моделі, призначеного для платформи IoT, за допомогою якого потрібно організувати канал з відповідними показниками QoS для виконання найсуворіших критеріїв якості надання послуг у перспективних мережах зв'язку п'ятого покоління;

Взаємодії відбуваються між такими пристроями:

- Інтернет Річ і платформа IoT;
- платформа IoT і сервіс провайдера, при цьому взаємодія відбувається по захищеному, секретному каналу зв'язку;
- платформа IoT і інша платформа IoT;
- додаток моделі на оркестраторі – додаток одного оператора зв'язку та іншого оператора зв'язку.

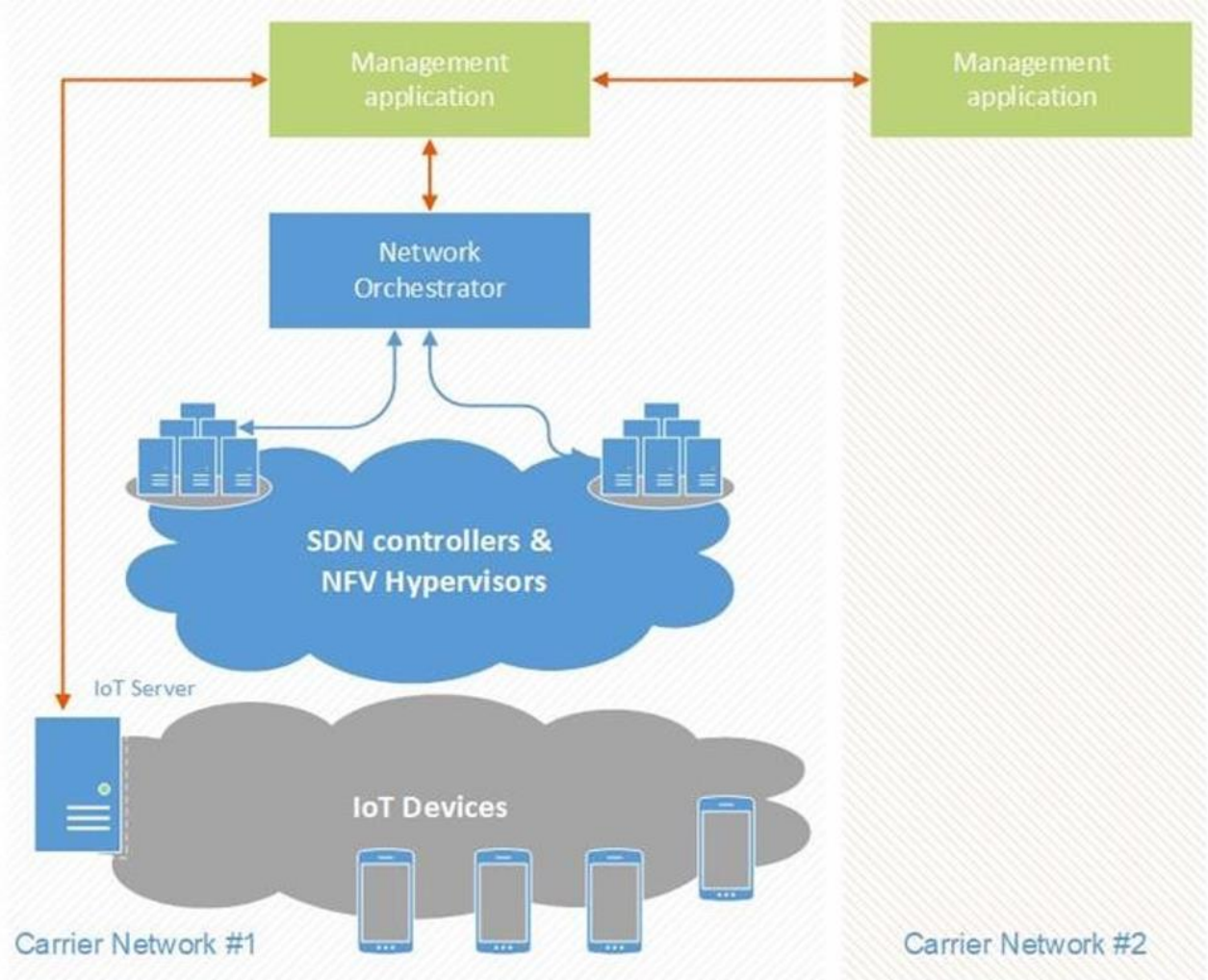

Рисунок 4.3 Взаємодії між елементами моделі контролю QoS

Організація елемента моделі (Framework's App), у вигляді додатку оркестратору дозволить реалізувати послуги «прозоро», по всій підконтрольній мережі оператора, при цьому на даному рівні реалізовані інші інформаційні системи оператора, в тому числі реалізують функціонал по білінгу, OSS/BSS (Operation Support System/Business Support System), моніторингу, аналізу даних і так далі, що дозволить легко інтегрувати даний функціонал при реалізації фреймворка.

4.2.3 Функціональні елементи моделі

IoT-сервер – це сервер, який може бути представлений в програмному або програмно-апаратному вигляді. Постачальник послуг надає низку послуг, розгорнутих на серверах, що знаходяться під його контролем, наприклад: зберігання даних з IoT-пристроїв і їх обробка, надання брокера MQTT (Message Queue Telemetry Transport), забезпечення встановлення зв'язку між IoT-пристроями.

Таким чином, IoT-сервер реалізує наступні основні функції:

- взаємодія з IoT-device. Взаємодія може бути організована на основі одного з протоколів Інтернету Речей (HTTP 2.0 (Hyper Text Transfer Protocol), CoAP (Constrained Application Protocol), MQTT, і т.д.);
- реалізація функціональності AAA для зареєстрованих IoT-devices.
- взаємодія з Management Application (MA). Взаємодія відбувається через API MA, при цьому вибір протоколу визначається реалізацією MA;
- взаємодія з іншим IoT-сервером. Дана взаємодія передбачається при наданні послуги, яка вимагає з'єднання «IoT-device - IoT-device», за участю IoT-сервера. У цьому випадку використовуваний протокол визначається реалізацією IoT-серверів.

Management Application (MA) – це сервер, який може бути представлений як в програмному, так і апаратно-програмному вигляді. Основні функції MA такі:

- перевірка можливості взаємодії з IoT-сервером. Для забезпечення якості послуг, що надаються, оператор послуг укладає угоду з оператором мережі, в результаті чого оператор мережі вносить IP-адреси сервіс-оператора до реєстру. Цей реєстр використовується оператором мережі для перевірки на етапі запиту на взаємодію від IoT-сервера. Також на даному етапі визначається ряд послуг, що надаються оператором мережі сервіс-оператору (наприклад, обмеження за параметрами QoS);

- взаємодія з IoT-сервером. Цей зв’язок відбувається через API МА за допомогою певного протоколу прикладного рівня. Програмний інтерфейс реалізує набір функцій, досяжність деяких визначається спочатку встановленою здатністю взаємодіяти з IoT-сервером під час процесу початкової перевірки. Основними функціями є прийняття запитів на обслуговування певного пристрою/групи пристроїв від IoT-сервера, а також сигнальний обмін з IoT-сервером в процесі супроводу IoT-пристроїв і надання відповідної якості;
- взаємодія з оркестратором мережі. Дана взаємодія відбувається через API оркестратор;
- взаємодія «МА-МА». Даний вид взаємодії можливий при необхідності встановлення з'єднання або з IoT-сервером, що знаходиться в мережі, підконтрольної іншому оркестратору (цього ж або іншого оператора мережі, наприклад, міжнародний роумінг), або при встановленні з'єднання «IoT- device - IoT-device», з урахуванням того, що один з пристроїв знаходиться в мережі, підконтрольної іншому оркестратору.

4.2.4 Організація контролю QoS для пріоритезації додатків

1. «Процес ідентифікації» - даний процес реалізує набір процесів, які спрямовані на виконання наступних трьох дій: ААА. При цьому даний процес відбувається тільки між Інтернет Річчю (віртуальною або фізичною) і сторонньою платформою IoT. При цьому в процесі аутентифікації відбувається порівняння запиту «Речі» при встановленні з'єднання і виділення їй обчислювальних ресурсів платформи згідно заздалегідь заведеній в базу даних інформації, що дозволяє зіставити дану річ і конкретного користувача, при цьому Інтернет Річ проходить процес аутентифікації згідно унікальному ідентифікатору. Процес авторизації дозволяє визначити ряд послуг, доступних цій речі, при цьому варто також врахувати, що в деяких випадках «Річ» може

реалізовувати певний набір функцій, характерних за своїми завданнями різному типу Речей, тобто, наприклад, Річ може реалізовувати функції Тактильного Інтернету , а також одну і/або декілька медичних функцій, відповідно з неоднорідністю характеристик Інтернет Речей можливо їх поділ на групи в моделі. В процесі accounting на сервері IoT ведеться облік використання обчислювальних ресурсів, виділених для конкретної Інтернет Речі, також ведеться логування процесів при їх можливому збої, в деяких випадках можлива також реалізація білінгу на стороні сервіс-оператора (власника IoT-платформи). Варто відзначити, що в даному процесі можуть бути реалізовані і інші функції, проте в кожному окремому випадку, набір функцій визначається сервіс провайдером (власником IoT-платформи).

2. «Iot group options» - формування сервером IoT і подальша передача параметрів групи Інтернету Речей і/або додавання до цієї групи Інтернет Речі з даними за необхідним QoS, протокол передачі даних який використовується (MQTT, CoAP і т.д.), рівень безпеки (потрібно додатково шифровані «канали» зв'язку) певної ділянки мережі.

3. «Configure VM for IoT-group» - після виконання вищестоящої операції, модель здійснює передачу запиту на оркестратор мережі відповідно до переданих даних в процесі «lot group options», після чого оркестратор здійснює конфігурацію віртуальної площадки (VM) для розгортання параметрів даної моделі під конкретну групу пристроїв IoT і/або проводить доповнення/очищення групи.

4. Create VM Space - створення віртуальної площадки для підсистеми групи IoT. Тобто, наприклад, створюється відокремлена «віртуальна площадка» для групи Інтернет Речей за певними критеріями обслуговування (різні сервіс оператори IoT, QoS, фінансове регулювання або білінг, DPI-вимоги та ін.). Або додаються до цієї групи нові зареєстровані пристрої IoT. Якщо розглянути за прикладом: сервіс-оператор Інтернету Речей «А» (власник IoT-платформи, що надає певний набір послуг і має контракт з оператором мережі зв'язку на підтримку роботи за цією моделлю) і, наприклад, рівноцінний по правам

доступу сервіс-оператор «B», що працюють в одному сегменті мережі, повинні бути логічно відокремлені в моделі.

5. Наступним кроком є процес «Set Framework», який включає в себе введення і налаштування параметрів мережі, QoS, протоколу IoT-групи. Слід зазначити, що на даному етапі можливі також і додаткові процеси («Set type, IoT method Identification»), які необхідні в умовах міжоператорських з'єднань. Як вже зазначалося вище, одним з взаємодій в даній моделі є взаємодія «Framework's App # 1 - Framework's App #N», при цьому дана взаємодія охоплює і інші процеси, такі як: «Create VM Space», «Set Framework».

6. Наступним кроком є «Set options for Monitoring Systems, OSS/BSS, Cellular Information Systems, b2b». Дані процеси спрямовані, при необхідності, на інтеграцію з системами моніторингу, OSS/BSS та іншими системами, необхідними оператору при наданні послуг клієнту. При цьому на рисунку 4.4 можна також помітити зворотний зв'язок між процесами, що має на увазі реалізацію повноцінного моніторингу послуги за допомогою моделі інформування моніторингових систем не тільки оператора, але і сервера IoT. Наприклад, в нештатній або будь-якій іншій ситуації, коли мережа, як обмежена в ресурсах система, не зможе далі забезпечити необхідну якість надання послуги, то на основі пріоритетів, які залежать як від типу додатка Інтернету Речі, так і від договору між оператором мережі та власником IoT-сервера, модель буде змушена відмовити в продовженні надання послуги, при тому, повідомити код помилки сервера. Зворотній зв'язок з сервером IoT дозволяє проводити не тільки аналіз та інформування систем один одного, але і прогноз збільшення кількості Інтернету Речей в кожному з секторів, їх тип і так далі, що в підсумку дозволить виробляти більш оптимальне бізнес-планування операторів мереж зв'язку, сервіс-операторів IoT і т.д.

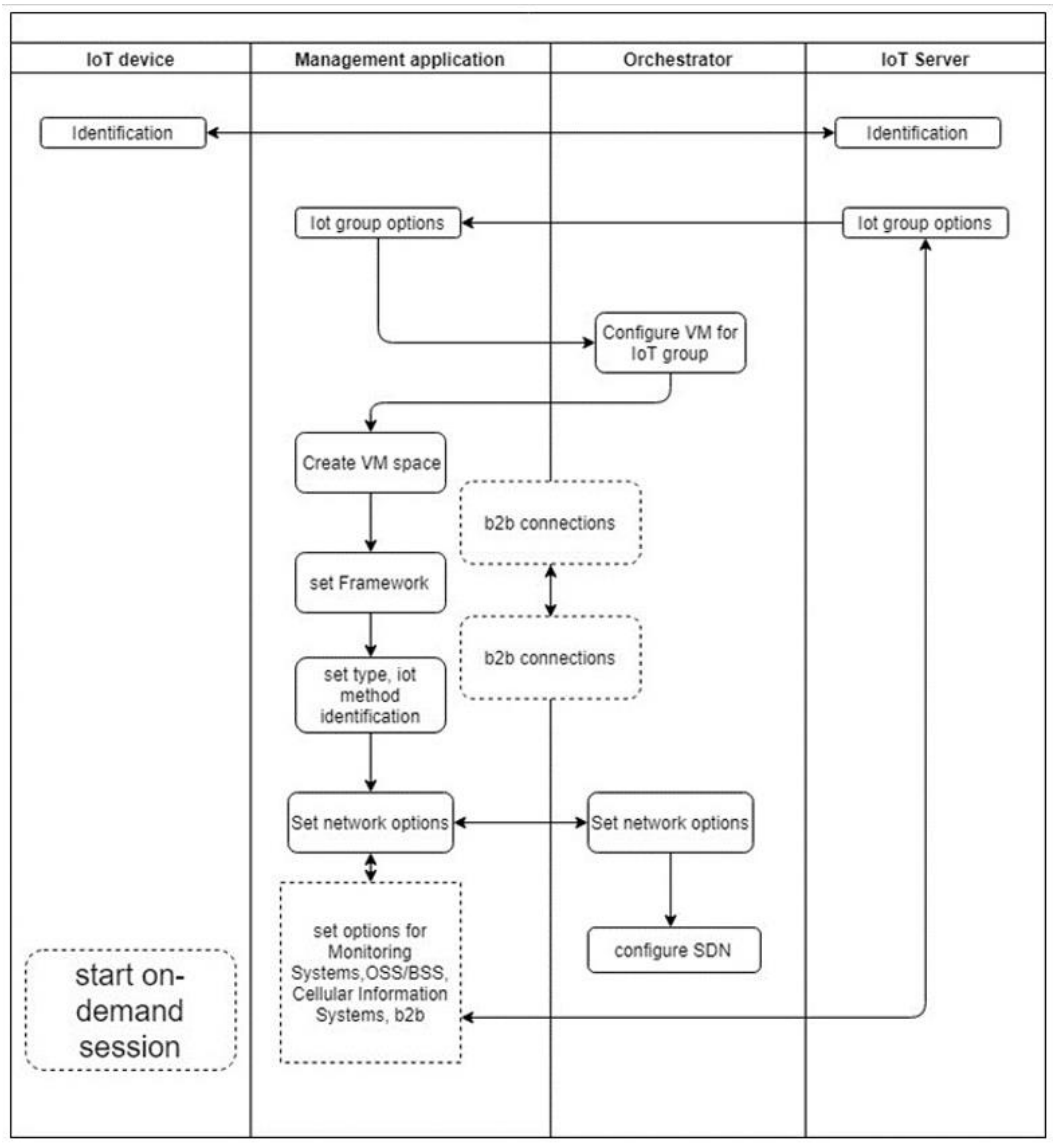

Рисунок 4.4 Процес взаємодії між основними елементами моделі

4.2.5 Моделювання сегмента запропонованої мережі

У цій частині оцінка продуктивності запропонованої моделі проводиться в надійному середовищі моделювання заснованому на Java і побудованому на основі CloudSim. Воно дозволяє створювати прикордонні хмари з різною кількістю віртуальних машин. Моделювання виконується на робочому комп'ютері з процесором Intel Core i5, частотою 3,07 ГГц і об'ємом пам'яті 8 ГБ. Результати моделювання ілюструються в таблиці 4.1.

Таблиця 4.1 – Параметри моделювання

Параметр	*Опис*	*Значення*
S	Число джерел	20
N_S	Кількість сегментів	6 (Додатки)
NVM	Число віртуальних машин	8
n	Розмір заголовку ідентифікатора	3 бита
W_{mmax}	Максимальне навантаження сервера	100 флопс/с
λ	Інтенсивність надходження	15 Мб/с
μ	Інтенсивність обслуговування сервера	8 Мб/с
RAM, HDD	RAM, Память	2048 Мб, 10 Гб

Створено прикордонний сервер з вісьмома віртуальними машинами. Прикордонний сервер MEC працює на основі моделі для прикордонних серверів Micro-cloud. Двадцять різних джерел створюють та генерують дані для шести різних програм, кожна програма трактується як окремий сегмент. Кожна програма резервує частину ресурсів на прикордонному обчислювальному сервері (MEC server). Ресурси для додатків вважаються неоднорідними, тому деяким програмам виділяється більше ресурсів, ніж іншим, на основі попередньо розподіленої схеми. Ця схема встановлюється на основі ймовірності запитів і потоку даних, виділених для кожної програми. Для нашої реалізації кількість віртуальних машин, виділених для кожної програми, включено в таблицю 4.2. Передбачається, що додатки 2 і 3 мають високий потік і, таким чином, виділяються обидві віртуальні машини.

Таблиця 4.2 Розподіл ресурсів для різних додатків

APP#1	APP#2	APP#3	APP#4	APP#5	APP#6
VM#7	VM#0 & VM#1	VM#6	VM#2&VM#3	VM#4	VM#5

Для оцінки продуктивності розглядаються два основних параметри продуктивності: затримка і ймовірність блокування. Ми вимірюємо кожен параметр для кожного додатка в двох випадках. Перший випадок являє запропоновану структуру, в якій ресурси виділяються для кожної програми. Другий випадок представляє альтернативний випадок, в якому сервера MEC служить без будь-яких класифікацій.

На рисунку 4.5 показано середню затримку для кожної програми в обох розглянутих випадках. Для всіх програм очевидно, що сегментація мережі забезпечує кращу продуктивність.

На рисунку 4.6 показує середню ймовірність блокування для кожного додатка в обох випадках. Імовірність блокування набагато нижча у випадку сегментації для всіх розглянутих додатків.

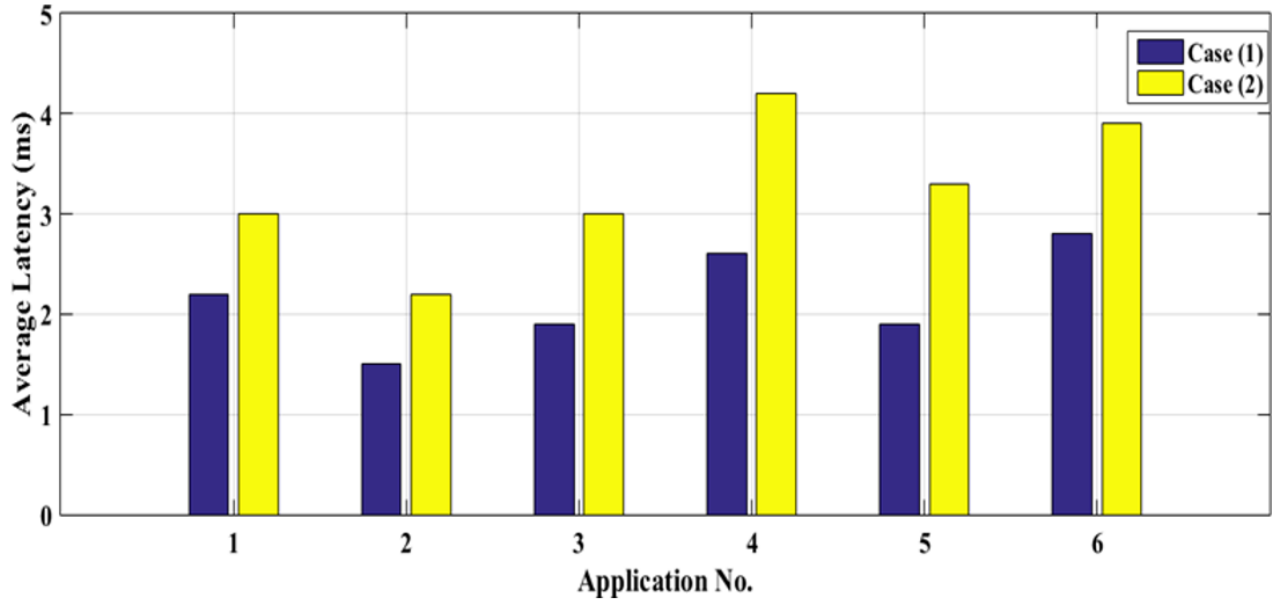

Рисунок 4.5 Середній час затримки для кожного додатка в розглянутих випадках

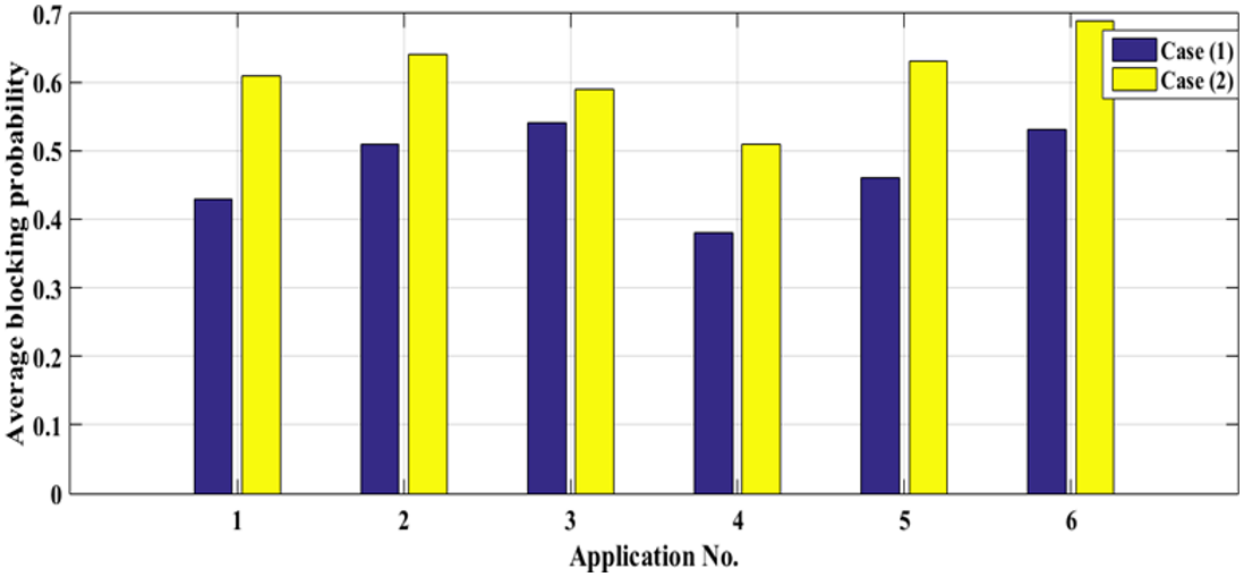

Рисунок 4.6 Середня ймовірність блокування для кожної програми в розглянутих випадках

Технологія SDN разом з NFV забезпечує гнучкість мережі, дозволяючи динамічно надати мережеві сегменти за потребою. Нові додатки можна легко розгорнути в мережі відповідно до вимог замовника. Сегментація мережі радикально спрощує процес представлення мережевих послуг у мережах 5G порівняно з традиційними моделями реалізації, дозволяючи динамічно розподіляти мережеві ресурси відповідно до вимог: пересилання трафіку з урахуванням вимог до обслуговування, управління трафіком з урахуванням фізичного стану каналів, об'єднання хмарних ресурсів та пропускної здатності мережі. Завдяки мережевій сегментації 5G здатна адаптуватися до зростаючих вимог до якості обслуговування, що дозволяє створювати нові бізнес-моделі на основі вимог користувачів / додатків.

4.3 Висновки до розділу 4

В розділі №4 проводиться розробка моделі ідентифікації та пріоритезації трафіку Інтернету речей на основі сегментації ресурсів.

Підсумком проведених в 3 розділі досліджень являється модель мережі оператора, яка традиційно розділена на 3 сегменти: рівень доступу, рівень агрегації, рівень ядра мережі. При цьому кожним з цих сегментів управляє ПКМ-контролер. За взаємодію між вище указаними сегментами відповідає оркестратор.

Також в рамках даного розділу було проведено імітаційне моделювання запропонованої моделі в середовищі CloudSim. В даному середовищі було створено прикордонний сервер з вісьмома віртуальними машинами. Двадцять різнорідних джерел створюють і генерують трафік для шести додатків, кожен додаток можна розглядати як окремий сегмент. Для оцінки продуктивності було виконано вимірювання затримки і ймовірності блокування. Кожен з цих параметрів вимірювався окремо для двох випадків: перший випадок являє собою сегментовану структуру, а другий звичайну мережу в якому сервера MEC служать без будь-яких класифікацій. На основі вимірювань видно, що мережева сегментація значно спрощує процес надання мережевих послуг в мережах 5G в порівнянні з традиційними моделями реалізації.

ВИСНОВКИ

Протягом останніх декількох років телекомунікаційна індустрія спостерігала тенденцію зростання кількості мереж фіксованого та мобільного зв’язку, наслідком цього являється гостра конкуренція між операторами та війна за абонентів. Це все привело до перенасичення ринку телекомунікацій, де користувачі отримують лише різні цінники на однакові послуги. Все вище сказане свідчило про необхідність прийняття рішень про шляхи подальшого розвитку сервіс провайдерів.

Ще кілька років тому багатьма провайдерами послуг було підняте питання про уніфікацію комунікацій як варіант вирішення цієї проблеми. Сьогодні це трансформувалось в питання про конвергенцію в телекомунікаціях. Тому вже зараз нерідко консолідуються компанії – провайдери різних типів послуг (фіксована і мобільна телефонія, мобільна телефонія і кабельне телебачення і т.п.). Однак в майбутньому такі процеси приведуть до появи операторів-гігантів, в яких буде потужна база користувачів мобільних та стаціонарних мереж, що поставить перед даними компаніями нову проблему – забезпечення конвергенції даних мереж між собою. Дана проблема викликала необхідність створення універсальної мережевої інфраструктури, такої як IP Multimedia Subsystem (IMS). Вона дозволяє забезпечити надання широкого спектру нових послуг усім користувачам мережі, незалежно від місця їхнього розташування, або ж термінального обладнання яке вони використовують, при цьому забезпечуючи найбільшу з можливих якість сервісу відносно кожного з абонентів.

Однак, впровадження концепції IMS не зменшить витрат оператора. Запуск будь-якого нового мережевого сервісу передбачає витрати на обслуговування, ремонт, заміну та купівлю нового обладнання, що вимагає місця в апаратних кімнатах, нових джерел живлення. Крім того, апаратні мережеві пристрої все швидше застарівають, не так фізично, скільки «морально», що вимагає все більш частих повторень циклу «закупівля -

проектування - інтеграція - розгортання». Стало ясно, що екстенсивний шлях розвитку операторських мереж на базі спеціалізованого обладнання є тупиковим. Потрібні нові підходи до розвитку бізнесу операторів і сервіс-провайдерів. Одним з таких підходів є віртуалізація мережевих функцій NFV, пов'язана з концепцією програмно-конфігурованих мереж SDN. Також поєднання цих технологій дозволяє реалізувати основну бізнес ідею мереж 5G – сегментування. Яка, в свою чергу, забезпечує гнучкість мережі, розбиваючи одну фізичну мережу на кілька шарів, кожен з яких має власні налаштування, адаптовані під певну послугу. Таким чином забезпечується ефективність і гнучкість майбутніх сервісів.

Проаналізувавши інформацію було з’ясовано, що SDN – це новий метод адміністрування і проектування комп'ютерних мереж. Ця технологія відокремлює площину управління мережею (Control plane), яка займається маршрутизацією трафіку, від площини передачі даних (Data plane), яка передає трафік згідно з, отриманими від площини управління, правилами. Також SDN «консолідує» площину управління, та з допомогою інтерфейсу прикладного програмування API з’явилась можливість централізованого управління мережею. Таким чином, введення нових послуг на мережі прискорюється і полегшується. В свою чергу NFV це технологія віртуалізації фізичних мережевих елементів телекомунікаційної мережі, коли мережеві функції виконуються програмними модулями, що працюють на стандартних серверах і віртуальних машинах в них. Ці програмні модулі можуть взаємодіяти між собою для надання послуг зв'язку, чим раніше займалися апаратні платформи. SDN та NFV, загалом, не залежать один від одного, хоча NFV в значній мірі доповнює SDN, також консолідування цих технологій допоможе розмити границі між вендорами, розв’язати проблему росту пакетного трафіку та напряму допоможе операторам раціоналізувати витрати на обслуговування і побудову мережі, що, в свою чергу, автоматично збільшить дохід.

В ході дослідження було з’ясовано що існує певний недолік властивий програмно-конфігурованим мережам. Відомо що основним завданням ПКМ-

контролера являється обробка вхідних запитів від комутаторів. Також не секрет що існує певний поріг навантаження при якому пристрій може вийти з ладу. Ці факти підводять нас до того, що використання єдиного централізованого контролера являє собою «вузьке» місце мережі, в разі його відмови мережа припинить функціонування. Одним з варіантів вирішення даної проблеми стала технологія розподілених контролерів, ідея яких полягає у використанні кількох контролерів на рівні управління, що дозволяє зменшити ризики відмови балансуючи навантаження між ними.

На основі цих даних було вирішено провести моделювання кластеру контролерів ПКМ, щоб оцінити ефективність їх функціонування. Експеримент по проектуванню імітаційної моделі фрагмента ПКМ було проведено в середовищі моделювання AnyLogic, дана модель описує процес надходження вхідних повідомлень на різні блоки обслуговування. Результати показали, що використання кластеру контролерів в 3 рази покращує можливості мережі в забезпеченні необхідної продуктивності.

Для реалізації завдань даної роботи був розроблений алгоритм кластеризації та пріоритезації для потоків трафіку мережі зв’язку. Метою алгоритму являється групування трафіку за певними ознаками для визначення приналежності потоку до певного кластеру (надалі сегменту). Що дозволяє налаштувати кластер під певну послугу та, наприклад, удосконалювати процес зняття плати.

Підсумком проведених досліджень являється модель мережі оператора зв’язку, яка традиційно розділена на 3 сегменти: рівень доступу, рівень агрегації, рівень ядра мережі. При цьому кожним з цих сегментів управляє ПКМ-контролер. За взаємодію між вище указаними сегментами відповідає оркестратор. В рамках розробки моделі було проведено імітаційне моделювання сегмента запропонованої мережі в середовищі CloudSim. При проведенні експериментального дослідження було отримано результати, які затверджують, що мережева сегментація значно спрощує процес надання

мережевих послуг в мережах 5G в порівнянні з традиційними моделями реалізації.

Отже, виходячи з усього вище сказаного, поставлені завдання, які були поставлені перед початком роботи над магістерською дисертацією, можна вважати виконаними.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

[1] Гольдштейн Б.С. Инфокоммуникационные сети и системы. – СПб.: БХВ-Петербург, 2019. – 207 с.: ил.

[2] Б. С. Гольдштейн, А. Е. Кучерявый. Сети связи пост-NGN – СПб.: БХВПетербург, 2014. —160 с.: ил.

[3] Гольдштейн Б.С., Гольдштейн А.Б. SOFTSWITCH – СПб.: БХВ – Санкт-Петербург, 2006.— 368 с.

[4] Blanco B., Technology pillars in the architecture of future 5G mobile networks: NFV, MEC and SDN / Blanco B., FajardoJ.O., Giannoulakis I., Kafetzakis E., Peng S., Pérez-Romero J., Trajkovska I., Khodashenas P.S., Goratti L., Paolino M., Sfakianakis E. //. 2017, Computer Standards & Interfaces.pp. 216–228.

[5] SDN: истоки создания и развития. //[Електронний ресурс] – режим доступу: https://shalaginov.com/2016/07/12/sdn-истоки-создания-и-развития/

[6] ONF Software-Defined Networking: The New Norm for Networks. //[Електронний ресурс] – режим доступу: https://www.opennetworking.org/images/stories/downloads/sdn-resources/whitepapers/wp-sdn-newnorm.pdf

[7] ONF OpenFlow Switch Specification. //[Електронний ресурс] – режим доступу: https://www.opennetworking.org/wp-content/uploads/2014/10/openflow-switchv1.5.1.pdf

[8] ETSI, Network Functions Virtualization (NFV)-ArchitecturalFramework.ETSIGSNFV002V1.2.1(Dez.2014). //[Електронний ресурс] – режим доступу: http://www.etsi.org/deliver/etsi_gs/nfv/001_099/002/01.02.01_60/gs_nfv002v010201p.pdf

[9] NGMN, Network Slicing Framework. //[Електронний ресурс] – режим доступу: https://www.ngmn.org/uploads/media/161010_NGMN_Network_Slicing_framework_v1.0.8.pdf.

[10] ITU-T Y.3011 Framework of network virtualization for future networks. //[Електронний ресурс] – режим доступу: https://www.itu.int/rec/T-REC-Y.3011-201201-I/en

[11] ITU-T IMT2020: Application of network softwarization to IMT-2020. //[Електронний ресурс] – режим доступу: http://www.itu.int/en/ITU-T/focusgroups/imt-2020/Pages/default.aspx.

[12] ITU-T Y.3150 (2018): High-level technical characteristics of network softwarization for IMT-2020. //[Електронний ресурс] – режим доступу: https://www.itu.int/rec/T-REC-Y.3150-201801-I/en

[13] ETSI GS NFV-EVE 012 (V3.1.1), Network Functions Virtualization (NFV) Release3; Evolution and Ecosystem; Report on Network Slicing Support with ETSI NFV Architecture Framework, December 2017. //[Електронний ресурс] – режим доступу: https://docbox.etsi.org/isg/nfv/open/Publications_pdf/Specs-Reports/NFV-EVE%20007v3.1.1%20%20GS%20%20NFVI%20Hw%20rqmts%20spec.pdf

[14] ETSI, Mobile Edge Computing A key technology towards 5G, ETSI White Paper, No. 11, September 2015. //[Електронний ресурс] – режим доступу: https://www.etsi.org/images/files/etsiwhitepapers/etsi_wp11_mec_a_key_technology_towards_5g.pdf

[15] ETSI. //[Електронний ресурс] – режим доступу: https://www.etsi.org/

[16] NGMN, description of network for service provider networks //[Електронний ресурс] – режим доступу: https://www.ngmn.org/fileadmin/user_upload/161010_NGMN_Network_Slicing_framework_v1.0.8.pdf

[17] NGMN, Network Slicing Framework. //[Електронний ресурс] – режим доступу: https://www.ngmn.org/uploads/media/161010_NGMN_Network_Slicing_framework_v1.0.8.pdf.

[18] NGMN. //[Електронний ресурс] – режим доступу: https://ngmn.org/

[19] 3GPP TR 22.830-030 (2018); Technical Specification Group Services and System Aspects; Feasibility Study on Business Role Models for Network Slicing. //[Електронний ресурс] – режим доступу: https://www.3gpp.org/ftp/Specs/archive/22_series/22.830/

[20] 3GPP TS 28.530-050; Management of network slicing in mobile networks; Concepts, use cases and requirements (Release 15), February 2018. //[Електронний ресурс] – режим доступу: https://www.3gpp.org/ftp/Specs/archive/28_series/28.530/

[21] 3GPP. //[Електронний ресурс] – режим доступу: https://www.3gpp.org

[22] 5G PPP, View on 5G Architecture, July 2016. //[Електронний ресурс] – режим доступу: https://5g-ppp.eu/wp-content/uploads/2017/07/5G-PPP-5G-Architecture-White-Paper-2-Summer-2017_For-Public-Consultation.pdf

[23] 5G PPP. //[Електронний ресурс] – режим доступу: https://5g-ppp.eu/

[24] IEEE. //[Електронний ресурс] – режим доступу: https://www.ieee.org/

[25] IETF, Interconnecting (or Stitching) Network Slice Subnets. //[Електронний ресурс] – режим доступу: https://tools.ietf.org/html/draft-defoy-coms-subnet- interconnection-03

[26] IETF, Software-Defined Networking (SDN): Layers and Architecture Terminology. //[Електронний ресурс] – режим доступу: https://tools.ietf.org/html/rfc7426

[27] GSMA, Introduction to network slicing. //[Електронний ресурс] – режим доступу: https://www.gsma.com/futurenetworks/wp-content/uploads/2017/11/GSMA-An-Introduction-to-Network-Slicing.pdf

[28] GSMA. //[Електронний ресурс] – режим доступу: https://www.gsma.com/futurenetworks/

[29] ONF OpenFlow Management and Configuration Protocol (OF-Config 1.1.1). //[Електронний ресурс] – режим доступу: http://opennetworking.wpengine.com/wp-content/uploads/2013/02/of-config-1-1-1.pdf

[30] ONF OpenFlow-enabled SDN and Network Functions Virtualization. //[Електронний ресурс] – режим доступу: https://opennetworking.org/wp-content/uploads/2013/05/sb-sdn-nvf-solution.pdf

[31] ONF TR-502 SDN architecture. //[Електронний ресурс] – режим доступу: http://opennetworking.wpengine.com/wp-content/uploads/2013/02/TR_SDN_ARCH_1.0_06062014.pdf

[32] ONF //[Електронний ресурс] – режим доступу: https://www.opennetworking.org/

[33] BBF SD-406 End-to-End Network Slicing. //[Електронний ресурс] – режим доступу: https://www.broadband-forum.org/5g

[34] Гимадинов Р. Ф., Кластеризация в мобильных сетях 5G. случай частичной мобильности / Гимадинов Р. Ф., Мутханна А. С., Кучерявый А. Е // Информационные технологии и телекоммуникации. – 2015. – № 2 (10). – С. 44–52. – ISSN: 2307–1303

[35] Mohammad S. M., Machine Learning for Internet of Things Data Analysis: A Survey / Mohammad S. M., Mohammadreza R., Mohammadamin B., Peyman A., Payam B., Amit P. S. // Digital Communications and Networks, Volume 4, Issue 3, August 2018, Pages 161–175.